COPYRIGHT

Luxembourg: Office des publications officielles
des Communautés européennes, 2001

ISBN 92-894-0433-7

© Communautés européennes, 2001

ÉDITION 2000

Manuel des concepts relatifs aux systèmes d'information sur l'occupation et l'utilisation des sols

5





De nombreuses autres informations sur l'Union européenne sont disponibles sur Internet
via le serveur Europa (http://europa.eu.int).

Une fiche bibliographique figure à la fin de l'ouvrage.

Luxembourg: Office des publications officielles des Communautés européennes, 2001

ISBN 92-894-0433-7

© Communautés européennes, 2001

Printed in Luxembourg

IMPRIMÉ SUR PAPIER BLANCHI SANS CHLORE

TABLE DES MATIÈRES

1. INTRODUCTION

1. INTRODUCTION

Ce manuel entend sensibiliser les experts et les néophytes aux différences techniques, méthodologiques et conceptuelles existant entre les systèmes d'information sur l'occupation et l'utilisation des sols.

Il devrait contribuer à une meilleure compréhension et à une validation plus objective des informations disponibles, tout en soulignant la nécessité d'une harmonisation des systèmes d'information au niveau européen.

Pour remplir ces objectifs, le manuel des concepts s'attache à :

- définir les termes de base (sol, occupation des sols, utilisation des sols) ;
- expliquer les concepts et principes fondamentaux relatifs aux éléments de base des systèmes d'information sur l'occupation et l'utilisation des sols (systèmes de classification, outils de collecte des données) ;
- expliquer, définir et décrire les termes techniques, souvent ignorés et pourtant indispensables au retour d'information ;
- identifier les besoins en données et les problèmes afférents.

Ce manuel apporte aux experts et aux néophytes une définition et un langage communs sur les différents éléments qui composent les systèmes d'information sur l'occupation et l'utilisation des sols.

2. DEFINITION DES TERMES GENERAUX

2. DEFINITION DES TERMES GENERAUX

2.1. Sol

Le terme de « sol » est largement utilisé, pourtant rares sont les définitions qui en sont données.

Le groupe de travail interdépartemental sur l'aménagement du territoire (IDWG-LUP) à la FAO (Organisation des Nations unies pour l'alimentation et l'agriculture) proposait en 1994 la définition suivante : *zone délimitée de la surface terrestre, qui englobe tous les attributs de la biosphère situés immédiatement au-dessus ou en dessous, y compris le climat près de la surface, le relief, l'hydrologie avec les lacs peu profonds, les rivières, les marécages et les marais, les couches sédimentaires de surface, les nappes d'eau souterraines et les réserves hydrogéologiques associées, la flore et la faune, l'habitat humain et les marques physiques de l'activité humaine présente et passée (aménagement de terrasses, réserves d'eau ou structures de drainage, routes, bâtiments, etc.).*

Cette définition soulève également la question du territoire à considérer : « tous les attributs de la biosphère situés immédiatement au-dessus ou en dessous ». Pour l'occupation des sols, la solution est simple : la zone de référence se trouve au-dessus de la surface (voir les définitions au chapitre suivant). Pour l'utilisation des sols, c'est plus complexe : d'un point de vue pragmatique, et compte tenu de l'importance économique de la diversité d'utilisation des zones « urbaines », la définition du sol devrait également englober les niveaux situés en dessous ou au-dessus. Les mines, les métros sous les zones urbaines, les nappes phréatiques, les champignonnières, etc., risquent donc de poser problème. Exemple : les zones d'extraction de pétrole ; faut-il prendre en compte tout le champ d'extraction (et, dans une certaine mesure, l'ensemble de la concession) ou seulement le puits ?

Afin de limiter les éventuels problèmes, nous proposons de restreindre son application aux cas raisonnables et d'accepter les usages « urbains » au-dessus et en dessous du niveau du sol (exemple des bâtiments comprenant des magasins en rez-de-chaussée, des appartements et des bureaux dans les étages, et des parkings en sous-sol). Les activités d'extraction de ressources naturelles (y compris le pompage des nappes phréatiques) restent le principal souci. Face à cette situation (et à la difficulté de recueillir des données pertinentes), nous suggérons de limiter la prise en compte de ces utilisations à leur impact physique au niveau du sol (puits de pétrole, entrée de champignonnière, etc.). Il est possible de généraliser ce principe à n'importe quel type d'utilisation.

Étant donné le contexte des systèmes d'information mondiaux sur l'occupation et l'utilisation des sols, et la difficulté de tracer des frontières précises entre la terre et l'eau (notamment dans les zones inondées), il est conseillé **d'étendre le concept de sol aux eaux intérieures et aux bancs découverts à marée basse**.

Cette proposition de définition doit être clairement séparée des concepts appliqués par les statisticiens pour délimiter un territoire à des fins statistiques. Eurostat a récemment proposé (EUROSTAT 1999) de recourir au concept statistique en vigueur, qui exclut les lacs, les rivières et les zones côtières. Cette approche se conçoit aisément dans le cadre, par exemple, des calculs des densités de population quand 17 % d'un pays tel que les Pays-Bas sont recouverts par les eaux. Le point de vue d'Eurostat procède de la nécessité de fournir des données statistiques harmonisées, le meilleur exemple étant le calcul des densités de population (les chiffres néerlandais sont radicalement différents si les zones inondées sont incluses dans la superficie totale du pays).

2.2. Objets

Si, dans de nombreux domaines scientifiques, la question des unités ou objets ne se pose pas, ce n'est pas toujours le cas pour ce qui est du sol : la signification d'un objet est problématique dans la mesure où la classification catégorielle d'une partie de la surface terrestre présuppose que la zone est clairement définie dans l'espace (DUHAMEL & VIDAL 1998). Les objets sont facilement identifiables si les espaces sont des superficies agricoles ou urbanisées, avec des frontières physiques. Or, ces dernières deviennent floues dans les environnements naturels ou semi-naturels. Les problèmes de délimitation sont aggravés dans les zones de transition.

Le *continuum* biophysique de l’environnement méditerranéen, par exemple, présente des transitions floues entre la forêt, le maquis et les prairies et pelouses sèches. Des problèmes de délimitation surgissent également là où sont utilisées des définitions catégorielles reposant sur des pourcentages d’occupation ou d’utilisation.

Objets mixtes

Il est possible de définir trois types de combinaisons :

- Les combinaisons spatiales « par juxtaposition », qui dépendent surtout de l’échelle et des unités d’observation. Les statisticiens appliquent la règle du *prorata* (exemple de cultures associées à des biotopes naturels). De nombreuses « nomenclatures » cherchent à résoudre le problème en créant des classes mixtes, qui nécessitent des légendes puisqu’elles ne respectent pas le principe de l’absence de chevauchement ;



Figure 2.1 : Juxtaposition d’occupation et d’utilisation

- Les combinaisons spatiales en 3-D sont créées « par juxtaposition » de plusieurs occupations ou utilisations. Quelle que soit la résolution de l’observation, la décision d’opérer une distinction entre les différentes occupations ou utilisations doit s’appuyer sur une règle. Une meilleure résolution n’a aucun impact sur l’observation des superpositions : les combinaisons spatiales tridimensionnelles sont indépendantes de l’échelle. Là encore, les statisticiens appliquent la règle du *prorata* (exemple de cultures associées à des surfaces boisées). Des règles spécifiques, telle que la prévalence, sont parfois adoptées : usages dominants pour les superficies urbanisées, notamment.



Figure 2.2 : Illustration de superposition

- Les combinaisons temporelles d'occupation ou d'utilisation dépendent surtout de la période d'observation *dt*. Dans le domaine agricole, les statisticiens appliquent la règle de la récolte principale, c'est-à-dire celle dont la valeur commerciale est la plus élevée. (Si cette règle n'est pas applicable, la culture qui occupe le plus longtemps la parcelle est prise en compte.)

2.3. Occupation et utilisation des sols

Les systèmes d'information mélangent souvent occupation et utilisation des sols : les végétations naturelles et semi-naturelles y sont décrites en termes d'occupation des sols, les superficies agricoles et urbanisées en termes d'utilisation des sols. Or, la distinction entre ces deux notions est fondamentale, même si elle est souvent ignorée ou oubliée. La confusion et l'ambiguïté conduisent à des problèmes d'ordre pratique, notamment quand des données présentées dans différentes dimensions doivent être mises en corrélation, comparées et/ou combinées.

L'occupation des sols est une description physique de l'espace – *l'occupation (bio)physique observée de la surface terrestre* (DI GREGORIO & JANSEN 1997), c'est-à-dire ce qui recouvre le sol. Elle distingue plusieurs catégories biophysiques – les zones de végétation (arbres, buissons, champs, pelouses), les sols nus (même s'il s'agit d'un manque de couverture), les surfaces dures (roches, bâtiments), les surfaces humides et les plans d'eau (nappes et cours d'eau, zones inondables). Cette description a des répercussions sur les systèmes de classification, de collecte des données et les systèmes d'information en général. L'occupation des sols est « observée », c'est-à-dire scrutée par différentes « sources d'observation » situées à plus ou moins grande distance de la surface terrestre : l'œil humain, les photographies aériennes, les sondes satellites.

L'utilisation des sols recoupe plusieurs approches, avec deux « écoles » principales. La dimension *fonctionnelle* renvoie à la description des zones selon leur finalité socio-économique : superficies à vocation résidentielle, industrielle ou commerciale, agricole ou forestière, destinées aux loisirs ou à la préservation, etc. Des liens avec l'occupation peuvent être établis ; il est possible de déduire l'utilisation d'un sol à partir de son occupation, et inversement. Mais les situations sont souvent compliquées et le lien n'est pas toujours évident. L'approche *séquentielle* a été surtout développée pour les statistiques agricoles. Elle englobe une série d'opérations humaines visant à tirer des produits et/ou des bénéfices des ressources du sol. Exemple : une séquence d'opérations telles que le labourage, l'ensemencement, les traitements herbicides et fertilisants, et la récolte (MÜCHER et al. 1993).

Contrairement à l'occupation, l'utilisation des sols n'est pas aisée à « observer ». À titre d'exemple, il est souvent difficile de déterminer si des étendues herbeuses sont destinées à un usage agricole. Les données transmises par la source d'observation ne sont pas toujours suffisantes et un complément d'information peut s'avérer nécessaire. Dans le cas d'un usage agricole, les agriculteurs peuvent confirmer la présence ou non de bétail, le pacage, etc. Il est également possible de se référer aux caractéristiques indiquant la présence ou l'absence de bétail. Dans l'approche FONCTIONNELLE, le lien avec l'occupation des sols peut apporter une aide. Avec l'approche SÉQUENTIELLE, il peut être nécessaire de recueillir des données plus complètes regroupant différents attributs, par exemple dans une approche multitemporelle. Dans la suite de ce document, l'utilisation des sols sera considérée comme FONCTIONNELLE.

2.3.1. Interrelation entre occupation et utilisation des sols

Une comparaison avec les méthodes de classification des marchandises, où les objets sont décrits en fonction de leur matériau et de leur finalité, peut s'avérer très utile.

Il est parfois possible de distinguer l'**aspect fonctionnel** de l'**aspect biophysique** (DUHAMEL & VIDAL 1998). Une parcelle couverte par un champ de blé peut raisonnablement être associée à un usage agricole. De la même façon, l'aspect fonctionnel peut renseigner sur l'aspect biophysique. Il est logique de penser qu'une zone de production forestière correspond à une classe biophysique de type « arbre ». Dans d'autres cas, cependant, une catégorie biophysique peut renvoyer à une multitude de catégories fonctionnelles. Une étendue herbeuse, par exemple, peut être une pelouse en milieu

urbain, une piste d'aéroport, une prairie ensemencée, une pâture maigre, un parcours de golf – ou même un toit d'église en Islande. À l'inverse, la même classe fonctionnelle peut recouvrir plusieurs catégories biophysiques : une zone résidentielle se compose de pelouses, de bâtiments, de routes goudronnées, d'arbres et de sols nus.

Des arguments méthodologiques et techniques plaident en faveur d'une séparation systématique des deux approches. Même s'il est difficile de la justifier lors de l'analyse des besoins des utilisateurs et des coûts imputables à l'acquisition, à l'exploitation et à la gestion des données issues de différentes approches, la nécessité de connaître les deux dimensions peut être illustrée par l'exemple ci-dessous (adapté de LUND 1998) :

Imaginons un système d'information qui couvre uniquement l'aspect « occupation des sols » à différentes dates d'observation :

	t1	**t2**	**t3**	**t4**
OCCUPATION DES SOLS	ARBRES (Châtaigniers)	ARBRES (Châtaigniers)	ARBRES (Châtaigniers)	ARBRES (Châtaigniers)

La séquence ARBRES>ARBRES>ARBRES>ARBRES suggère une interprétation simple : aucun changement sur la période t1-t4 considérée.

Prenons la même zone, observée par un système d'information qui s'intéresse uniquement à l'utilisation des sols :

	t1	**t2**	**t3**	**t4**
UTILISATION DES SOLS	Vocation forestière (Bois de construction)	Vocation agricole (Production de châtaignes)	Vocation agricole (Pacage)	Vocation forestière

La séquence FORESTIÈRE>AGRICOLE>AGRICOLE>FORESTIÈRE pourrait facilement être interprétée comme le signe d'une déforestation entre les périodes t1 et t2, puis d'une reforestation entre t3 et t4. Avec un niveau de détail supplémentaire, la séquence « Bois de construction>Production de châtaignes>Pacage>Vocation forestière » donne une image plus précise de l'occupation des sols : entre t1 et t2, il est évident que le bois de construction et la production de châtaignes correspondent à une couverture boisée ; ce n'est pas aussi apparent entre t2 et t3 (production de chataignes>pacage).

	t1	**t2**	**t3**	**t4**
OCCUPATION DES SOLS	ARBRES (Châtaigniers)	ARBRES (Châtaigniers)	ARBRES (Châtaigniers)	ARBRES (Châtaigniers)
UTILISATION DES SOLS	Vocation forestière (Bois de construction)	Vocation agricole (Production de châtaignes)	Vocation agricole (Pacage)	Vocation forestière

En conclusion, la collecte simultanée de données sur l'occupation et l'utilisation des sols évite les erreurs d'interprétation ou de déduction et fournit des informations précieuses, exploitables dans de nombreuses applications.

2.4. Évolution de l'occupation et l'utilisation des sols

Les classes d'occupation et d'utilisation des sols sont des unités analytiques qui permettent d'établir un premier lien quantitatif entre les activités humaines, leurs répercussions sur l'environnement, et la dimension géographique (spatiale). Les informations concernant l'évolution de l'occupation/utilisation des sols intègrent la dimension temporelle.

Ces données présentent un intérêt majeur pour les politiciens – qui évaluent les décisions touchant à la terre – et pour la communauté scientifique – qui découvre les causes sous-jacentes et les conséquences.

En règle générale, les évolutions de l'occupation et l'utilisation des sols sont réparties en deux grandes catégories : la conversion et la modification (STOTT, A. & HAINES-YOUNG, R. 1996 ; ALUN, J. & CLARK, J. 1997 ; BAULIES, X.I. & SZEJWACH, G. 1997) :

- **La conversion** fait référence au passage de l'une des classes d'occupation ou d'utilisation à une autre (exemple : de la forêt aux prairies et pelouses).
- **La modification** représente une évolution à l'intérieur d'une classe (exemple : d'une zone de culture en sec à une culture irriguée) suite à des changements affectant ses attributs physiques ou fonctionnels.

Les informations purement relatives à l'occupation et à l'utilisation des sols prennent une dimension beaucoup plus intéressante à travers l'analyse, l'identification et la description du processus en cours.

Les évolutions permettent de déduire certains processus, qui peuvent aussi servir de simples indicateurs, parmi lesquels :

Intensification

Flux représentant le passage d'une faible intensité d'occupation ou d'utilisation des sols à une plus forte (exemple : d'un milieu semi-naturel à une terre arable).

Extensification

Flux représentant le passage d'une forte intensité d'occupation ou d'utilisation des sols à une plus faible (exemple : d'une pâture aménagée à un milieu semi-naturel).

Boisement

Flux représentant la plantation ou la régénération naturelle d'arbres.

Déboisement

Flux représentant l'abattage des arbres.

Développement

Flux impliquant la transformation d'un espace ouvert en un milieu à vocation urbaine, industrielle ou de transport.

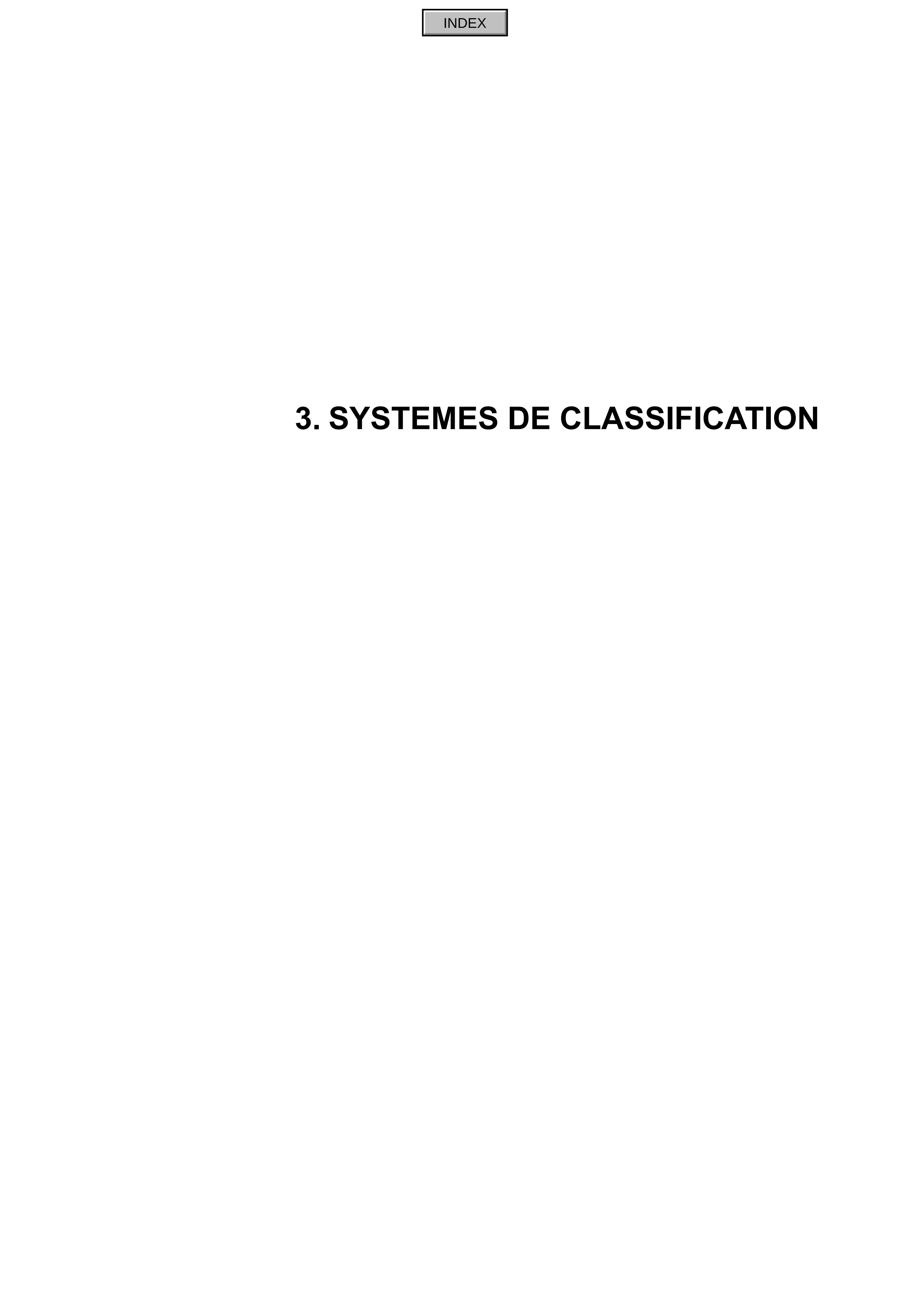
Réhabilitation

Flux impliquant la création d'un espace ouvert dans des zones précédemment développées (exemple : réhabilitation d'une exploitation minière).

La classification de ces processus peut être encore affinée en fonction de leur évolution (conversion ou modification), leur amplitude (portée du changement) et de leur rythme (vitesse d'évolution).

Notez que la valeur ajoutée des évolutions de l’occupation et l’utilisation des sols, ainsi que l’acquisition des processus et indicateurs sous-jacents, dépendent surtout des caractéristiques thématiques et spatiales associées à ces informations. Nous avons indiqué que les données sur l’occupation et l’utilisation des sols fournissaient en général des informations quantitatives (classes d’occupation/utilisation, estimations de la superficie). Des précisions concernant les caractéristiques qualitatives (pratique agricole, valeur écologique, etc.) sont nécessaires pour bien décrire les processus en action.

Cette approche comprend également l’analyse des éléments moteurs qui sous-tendent les évolutions de l’occupation et l’utilisation des sols, tels que les forces économiques, politiques, sociales et environnementales. Leur nombre et leurs interactions rendent cette analyse extrêmement complexe.



3. SYSTEMES DE CLASSIFICATION

3. SYSTEMES DE CLASSIFICATION

3.1. Définitions

De nombreux termes aux acceptions diverses sont utilisés en classification, en nomenclature et en taxinomie. Employés souvent de manière interchangeable, ils sont une source d'ambiguïté pour les lecteurs et ceux qui utilisent ces informations dans un domaine particulier. Leur signification est souvent issue des sciences naturelles, telles que la biologie, la botanique et la zoologie.

3.1.1. Classifications

La signification du terme de « classification », utilisé par les scientifiques du monde entier, remonte historiquement à la classification des organismes, d'après la systématique aristotélicienne. Ce terme est cependant ambigu dans la mesure où il recouvre deux acceptions principales (SUTCLIFFE, 1993) :

- L'établissement d'une **classification de tous les objets** dans l'univers de discours U (répondant à l'instruction : classifier les objets) ; le résultat est l'organisation de sous-classes dans U.
- L'utilisation de la classification ainsi établie pour décider de l'appartenance d'objets spécifiques (répondant à l'instruction : classifier l'objet X). Ce processus d'**identification** correspond à ce que les spécialistes en télédétection nomment classification.

La définition standard, largement répandue, de la classification de (SOKAL 1974) adaptée de (SIMPSON 1961) fait référence à *« l'ordonnancement ou l'arrangement d'organismes dans des groupes ou des ensembles en fonction de leurs relations »*. Cette définition a pris un sens plus général, dans lequel le terme *objets* remplace celui d'*organismes.* Elle englobe le classement de tous les objets, mais ne couvre expressément ni l'identification ni la désignation des groupes et les règles régissant l'usage de ces noms. Par ailleurs, le terme de relations, tel qu'il est employé en biologie, recoupe deux types de relations : les affinités, en termes de similitudes entre les caractéristiques des organismes (relations phénétiques), et les relations par les ancêtres (relations phylétiques), qui n'ont pas d'équivalent dans notre domaine de recherche.

3.1.2. Nomenclatures

Une nomenclature est une liste de catégories qui récapitule les informations sous une forme très synthétique tout en essayant de conserver le maximum de contenu. En principe, elle couvre un domaine d'intérêt particulier (l'univers de discours de SUTCLIFFE). En biologie, les définitions des nomenclatures renvoient à l'aspect légaliste de la taxinomie, à savoir la désignation des groupes et des organismes, ainsi que les règles régissant l'usage de ces noms.

3.1.3. Légendes

Les légendes sont fréquemment confondues avec les systèmes de classification et les nomenclatures. En principe, une légende est l'application d'une nomenclature dans un but spécifique : l'établissement de cartes thématiques avec une échelle et des unités cartographiques appropriées. Si la nomenclature est exhaustive et couvre l'intégralité de « l'univers de discours » ou du domaine, la légende en constitue un sous-ensemble. Certaines classes de la nomenclature peuvent être utilisées telles qu'elles, d'autres omises ou combinées, associées à des catégories composites (classes

mixtes) en respectant des contraintes minimales de délinéation à des fins de lisibilité. Si les systèmes de classification sont, et doivent être, indépendants de l'échelle et de la représentation cartographique, ce n'est pas le cas des légendes. Dans une certaine mesure, les systèmes de classification séparent les types d'occupation ou d'utilisation des sols ; les légendes, en revanche, divisent le territoire en fonction des attributs ou variables d'occupation et/ou d'utilisation. Par exemple, le projet CORINE-Land Cover propose une « nomenclature » qui est en réalité une légende, dans laquelle notamment certaines classes se recoupent dans des classes mixtes.

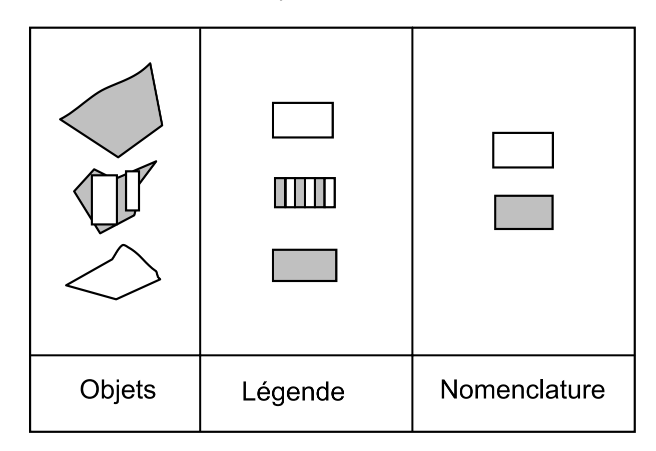


Figure 3.1 : Objets, légendes et nomenclatures

3.1.4. Vers des systèmes de classification

Étant donné les chevauchements de définitions et la nécessité d'élaborer des dispositifs de base pour structurer l'information et simplifier la communication et les échanges dans une discipline donnée (l'occupation et l'utilisation des sols, en l'occurrence), il est nécessaire d'identifier les fonctions indispensables à un « système de classification » pour satisfaire les exigences des scientifiques et des utilisateurs. Trois fonctions principales sont proposées (tirées en majorité de SUTCLIFFE 1993) :

Classification = Affectation de tous les objets dans une hiérarchie de catégories imbriquées, organisées de façon à montrer leurs interrelations.

Nomenclature = Désignation et description des groupes d'objets. Le résultat est une liste de noms et de descriptions mises en corrélation et présentées en général d'après la structure de la classification établie.

Identification = Affectation d'un individu à un groupe préalablement classifié et nommé. En d'autres termes, utilisation de la base de classification établie pour déterminer l'appartenance de chaque objet.

Il est recommandé d'appliquer le concept de **système de classification** englobant :

- la délimitation d'un univers de discours (U), en l'occurrence le domaine de l'occupation et de l'utilisation des sols (DOMAINE) ;
- l'établissement d'une classification de tous les objets appartenant à l'univers de discours (U), c'est-à-dire l'organisation des sous-classes de U selon une hiérarchie de catégories imbriquées, organisées de façon à montrer leurs interrelations (CLASSIFICATION) ;
- un système de désignation et de description des groupes associés à la structure de classification établie (NOMENCLATURE) ;
- des procédures d'affectation d'un objet à un et un seul groupe préalablement classifié et nommé (IDENTIFICATION).

3.2. Systèmes de classification

3.2.1. Une approximation de la réalité

Les systèmes de classification sont des outils qui décrivent certains aspects du monde réel. Les catégories sélectionnées n'en représentent pas une division unidimensionnelle, mais multidimensionnelle. La partition du monde réel met en exergue des aspects spécifiques de la réalité : une même réalité peut être reflétée dans plusieurs classifications (DUPRAT 1972).

En règle générale, les systèmes de classification se présentent sous la forme d'une arborescence, c'est-à-dire d'une hiérarchie. Un système hiérarchique organise les objets en série de groupes auxquels sont attribués différents niveaux de rang. Les groupes d'objets sont définis par l'attribution de caractéristiques communes aux membres d'un groupe, qui les distinguent des autres. Chaque division signifie que des caractéristiques « objectives » sont prises en compte et implique un choix conscient. Certains utilisateurs peuvent souhaiter scinder leur domaine d'intérêt à un niveau donné, selon des critères différents. Il en résulte de sérieuses difficultés lors des comparaisons : une classe considérée comme unique dans un système de classification peut correspondre à deux catégories ou plus dans un autre système reposant sur des principes distincts. Impossible, par exemple, de concilier avec exactitude une classification d'arbres qui distingue les subdivisions *à larges feuilles* et *à feuilles aciculaires*, et une classification qui n'identifie que les espèces *à feuilles caduques* et *à feuilles persistantes* (WYATT 1997).

Il convient également de noter qu'une catégorie de classification peut être homogène par rapport à une caractéristique (classe monothétique), ou deux, ou… aucune (suivant le concept des classes polythétiques proposé par les biologistes). Si l'agrégation dépasse un certain seuil, les catégories ne renvoient plus un reflet exact des entités significatives : c'est le cas d'un agrégat qui mélangerait, par exemple, des superficies agricoles et urbanisées au sein d'un système de classification.

Un système de classification repose donc sur une structure et un classement associés à un système de valeurs, qui révèle une intention. La finalité de la classification détermine nécessairement sa structure et son contenu. C'est la raison pour laquelle chaque utilisateur établit en général sa propre classification, adaptée à ses besoins spécifiques : or, l'élaboration spontanée de classifications conduit inévitablement à des incompatibilités.

3.2.2. Un compromis

Un système de classification devrait procéder d'un dialogue permanent entre :

- **une approche systématique** imposant une structure de l'information selon des principes logiques (exhaustivité, absence de chevauchement, définitions univoques des classes, règles régissant la représentation des objets dans la classification) ;
- **une approche pragmatique** prenant en compte les besoins de l'utilisateur et les informations existantes ;
- **une approche contextuelle** intégrant les contraintes propres au domaine d'étude. Dans le cadre de l'occupation et de l'utilisation des sols, certaines contraintes sont inhérentes à la dimension géographique de l'information.

3.3.3. Quelques propriétés

Cohérence spatiale

Le principe de ***cohérence spatiale*** implique des systèmes de classification qui permettent d'obtenir des résultats compatibles entre les différents sites, régions ou pays de la zone géographique considérée.

Cohérence temporelle

Les types d'occupation ou d'utilisation des sols devraient être consignés au moment précis de l'observation (par l'observateur, l'enquêteur ou la sonde satellite). Le système de classification ne doit donc pas tenir compte des situations passées ou futures (exemple : un programme de chantiers, qui reflète une intention d'utilisation). Les résultats doivent être considérés comme des stocks, et non comme des flux (ces derniers sont mesurés en comparant deux stocks). La ***cohérence temporelle*** repose sur ce principe. Toutefois, les zones où des changements sont intervenus sur une courte période doivent faire l'objet d'une attention particulière. À ce titre, la nomenclature de la Commission économique pour l'Europe des Nations Unies (CEE-NU) est un mauvais exemple, puisqu'à la catégorie 3.9.4 (*Terrains destinés à la construction*) correspond la définition « terrains désignés dans les plans publics d'occupation des sols comme des terrains constructibles, mais dont les travaux de construction n'ont pas encore démarré ». Le sens en est discutable : la construction n'a-t-elle pas commencé ou bien les données n'ont-elles pas été mises à jour depuis plusieurs années ?

Compatibilité avec les systèmes de classification existants

Un système de classification revêt un intérêt inestimable pour la compilation des statistiques internationales ou mondiales sur l'occupation ou l'utilisation des sols. La création d'un tel système à vocation générale se justifie uniquement s'il garantit un assez haut degré de compatibilité avec les systèmes d'information déjà en place. Il faut donc veiller avant tout à ce qu'il soit aussi ***compatible*** que possible ***avec les principaux systèmes existants***, afin de pouvoir tirer des conclusions pertinentes à partir des données issues de diverses sources fiables. À titre d'exemple, un système de classification sur l'utilisation des sols devrait être lié autant que possible à des classifications socio-économiques, puisque son domaine d'étude comporte une dimension socio-économique. Des « passerelles » doivent être prévues pour permettre ce type de comparaison.

Indépendance vis-à-vis des outils de collecte et de traitement des données

En théorie, le système de classification doit être établi sans tenir compte des ressources disponibles pour collecter l'information. Plus précisément, il devrait être autant que possible indépendant des échelles de restitution cartographique. Or, la pratique montre qu'il est difficile d'élaborer un système totalement affranchi de la méthode d'observation appliquée. Les classifications ont souvent été adaptées aux outils d'observation. Dans le domaine de l'occupation des sols, les systèmes sont conçus pour exploiter des outils spécifiques, tels que la photographie aérienne ou l'imagerie satellitaire. Quand les outils changent, le système de classification est incapable de s'adapter et des problèmes de continuité temporelle des informations peuvent survenir. Il arrive fréquemment qu'il faille trouver un compromis entre les outils disponibles, d'une part, et les besoins des utilisateurs, d'autre part (qui les uns comme les autres évoluent).

Exhaustivité

Chaque système de classification renvoie à un « segment » de la réalité (l'univers de discours), qui doit être décrit de manière exhaustive : toutes les « entités » physiques doivent figurer dans une classification de marchandises, tous les types d'occupations des sols dans une classification des occupations de sol et tous les paysages dans une classification de paysages. Selon la définition formelle de l'exhaustivité (BUNGE 1983) :

- il doit exister une classe pour tout objet à répertorier ;
- la réunion de toutes les classes de premier niveau doit refléter la collecte de données initiale.

Absence de chevauchement

Les classes doivent être exhaustives et ne pas se recouper. L'absence de chevauchement conditionne l'usage pertinent d'un système de classification : sinon, l'utilisateur est contraint de choisir entre plusieurs classes possibles. D'un point de vue formel (BUNGE 1983), l'absence de chevauchement signifie que :

- tout objet peut être répertorié dans une classe et une seule.

Dans le domaine de l'occupation et de l'utilisation des sols, ce principe peut être désigné sous le terme de ***cohérence sémantique*** (EUROSTAT 1992). L'une des principales conséquences est l'exclusion systématique des classes mixtes.

Règles d'identification

Les règles établies doivent être appliquées pour identifier et classifier les objets dans les catégories appropriées, en fonction de critères clairement définis. Il faut régler notamment la question des chevauchements et des panachages, ainsi que le problème des pièces détachées et des accessoires lors des classifications de marchandises.

Règles d'appellation

Le texte des *en-têtes* doit refléter le **contenu** et prendre en compte les caractéristiques du concept et les relations avec d'autres concepts :

- en général, il ne peut donner une description exhaustive, mais simplement une approximation, et doit refléter les caractéristiques du concept ;
- il doit être concis et abrégé (sinon l'utilisateur utilisera lui-même une abréviation), tout en restant compréhensible ;
- il doit être explicite (la définition d'une terre arable ne peut pas être « terre arable »).

Il doit également être en phase avec les connaissances et les besoins des utilisateurs.

Les *définitions* sont des descriptions verbales d'une notion à partir de concepts connus. La définition par **intension** (compréhension) est privilégiée : elle comprend des caractères restrictifs qui permettent de distinguer la classe des autres classes situées au même niveau dans la hiérarchie. Elle s'oppose à la définition par **extension** (générique) qui consiste à essayer d'énumérer tous les objets. Bien évidemment, il est utile de combiner les définitions par intention et par extension. Les **prototypes** ou types de catégories (spécimens désignés dans la terminologie biologique) peuvent aussi être utiles : les photographies de l'occupation des sols, par exemple, sont unanimement reconnues comme un prototype ou un représentant pertinent de la classe. Le guide technique du projet Corine Land Cover (1990) fournit de bons exemples.

Les *notes explicatives* donnent des détails supplémentaires sur les objets contenus dans une catégorie. C'est parfois la manière la plus simple de classifier, puisqu'elles indiquent les inclusions, les exclusions, les définitions et les décisions à prendre. Quand un objet est exclu d'une catégorie, il convient de préciser son appartenance. Après avoir identifié les « noyaux », il faut s'intéresser aux délimitations et répertorier les **conditions de limite** incluses ou exclues pour chaque catégorie. En cas d'exclusion, des références croisées doivent être établies vers la catégorie ad hoc.

Index

Un index alphabétique, constitué après le système de classification, est un outil très utile pour déterminer l'emplacement d'un objet dans la classification (IDENTIFICATION). Il est, par exemple, important de connaître à quelle classe appartient l'objet identifié comme un barrage : c'est le rôle de l'index, qui énumère tous les objets classifiés.

Principes de codification

Des codes sont associés aux images. En principe, ils sont structurés si la nomenclature est hiérarchique. Le principe de base prévoit que le code de niveau inférieur répète le code de niveau supérieur. Tous les symboles de codification sont normalement admis (chiffes, caractères, autres signes). L'une des questions importantes à régler est la signification du zéro (0) ; en règle générale, il doit être utilisé quand il n'existe aucune subdivision à un niveau donné.

Règles d'inclusion de nouveaux objets

Quels que soient les efforts déployés pour garantir l'exhaustivité du système de classification, il est inévitable que de nouveaux objets soient inclus par la suite. Les règles de classification associées à la nomenclature doivent alors être révisées, puis assignées à tous les objets concernés par la modification, y compris ceux qui avaient déjà été classifiés. La « jurisprudence » doit être examinée avec soin : il convient de s'assurer que l'objet à inclure est bien nouveau (d'où l'importance de l'index !) et modifier en conséquence le texte de la nomenclature.

Agrégats

Il est difficile d'avoir un texte, notamment la dénomination de la classe, qui reflète le contenu dans la mesure où le processus d'agrégation fait disparaître des informations. Pour assurer la pertinence des questions/réponses, entre autres, un certain niveau de détail est requis : dans les niveaux supérieurs d'agrégation, les catégories perdent leurs propriétés descriptives et deviennent de simples jalons documentaires (DUPRAT 1972). Même s'ils sont acceptés dans le langage commun, les agrégats sont souvent des unités à la définition imprécise, sans caractéristiques propres.

Hétérogénéité des champs sémantiques

En dehors de la problématique pure de la traduction, les champs sémantiques ne présentent pas toujours une correspondance totale et univoque. Voici l'exemple donné par (ECO 1988), adapté de (HJELMSLEV 1957) : le champ sémantique de forêt-bois-arbre montre l'absence de correspondance directe entre les unités lexicales de différentes langues (français, allemand, danois, italien et anglais). Selon les langues, la valeur sémantique est différente : la partition du *continuum* « forêt-bois-arbre » génère deux mots en danois, trois en français et en allemand, et quatre en italien et en anglais (sans distinction claire entre « timber » et « wood »). Voilà qui ne simplifie pas l'harmonisation des concepts.

<table>
<tr><th>Français</th><th>Allemand</th><th>Danois</th><th>Italien</th><th>Anglais</th></tr>
<tr><td>Arbre</td><td>Baum</td><td rowspan="2">Trae</td><td>Albero</td><td>Tree</td></tr>
<tr><td rowspan="3">Bois</td><td rowspan="2">Holz</td><td rowspan="2">Legno</td><td rowspan="2">Timber</td></tr>
<tr><td rowspan="3">Skov</td></tr>
<tr><td rowspan="2">Wald</td><td>Bosco</td><td>Wood</td></tr>
<tr><td>Forêt</td><td>Foresta</td><td>Forest</td></tr>
</table>

Figure 3.2 : Hétérogénéité des champs sémantiques (ECO 1988)

3.3. Types de système de classification

3.3.1. Les systèmes hiérarchiques

La présentation hiérarchique des nomenclatures est une construction artificielle ; dans la réalité, les objets sont liés par un réseau complexe de relations, similitudes, affinités ou corrélations. Afin de transformer ce réseau en arborescence, il est nécessaire de choisir les liens ou les relations perçus, dans un certain contexte, comme importants ou significatifs.

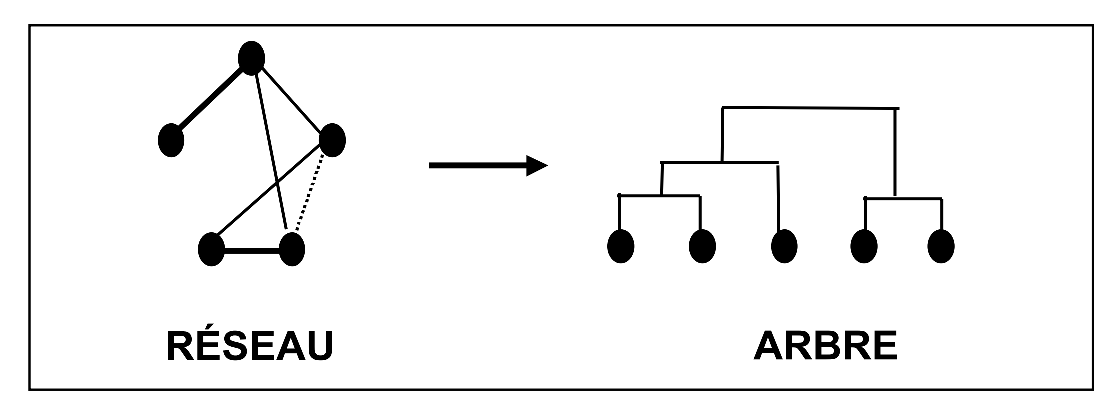


Figure 3.3 : Illustration des systèmes de réseau et d'arborescence

Il existe plusieurs types de liens ou de relations (adaptés de WÜSTER 1971) :

- liens logiques (industrie : chimie, transformation du fer, industrie automobile…) ;

- liens ontologiques = tout et partie (zone résidentielle - habitation) ;
- liens de coordination (maïs - blé) ;
- liens matériau-produit (vigne - vin) ;
- liens temporels (sol nu - récolte).

Ces liens et relations peuvent être combinés et considérés sous un angle plus large ou plus étroit. Dans une certaine mesure, les liens entre les occupations de sol sont surtout ontologiques (tout et partie : forêt > peuplement > arbre). À l'inverse, les liens relatifs à l'utilisation sont essentiellement logiques (industrie : chimie, transformation du fer, etc.).

Groupage

Des catégories sont constituées par regroupement d'objets « similaires ». Cette méthode implique de comparer un grand nombre de caractéristiques d'un objet aux mêmes caractéristiques d'autres objets. Ceux qui présentent de nombreuses propriétés communes sont regroupés et une nomenclature est ensuite établie à partir de cette répartition. Il existe plusieurs méthodes (analyse factorielle, classification ascendante hiérarchique, groupage dynamique), qui sont largement employées en botanique, en zoologie, en taxinomie linguistique sous le terme de taxinomie numérique (voir SNEATH & SOKAL 1973). Elles peuvent être utiles s'il existe suffisamment d'informations sur les unités d'observation. Or, c'est rarement le cas avec l'occupation et l'utilisation des sols.

Arborescence descendante

Cette solution apparaît comme la plus évidente et de nombreux systèmes de classification sont établis sur cette approche *a priori* : le champ d'étude est divisé en catégories et sous-catégories en fonction des objectifs poursuivis. Cette méthode présente des inconvénients majeurs : l'arborescence est une structure rigide, difficile à modifier sans altérer la structure initiale de l'information. La seule possibilité de modification implique la création de niveaux plus détaillés dans les catégories existantes. Quand l'arbre n'est plus adapté aux besoins, l'ajout de niveaux supplémentaires ne résout pas le problème. Autres inconvénients : le déséquilibre des secteurs et l'importance exagérée des concepts agrégés issus du premier découpage. La Commission économique pour l'Europe des Nations Unies (CEE-NU) a ainsi développé un système de description de l'occupation et de l'utilisation des sols. Il repose sur une nomenclature hiérarchique à trois niveaux, les deux premiers étant considérés comme fondamentaux, le troisième comme facultatif. Le premier niveau s'intéresse à l'occupation des sols et les deux autres, aux aspects fonctionnels. Cette nomenclature est assez bien adaptée aux climats tempérés et nordiques, mais s'avère extrêmement difficile à utiliser pour décrire les paysages de l'Europe du Sud. À titre d'exemple, la catégorie 5 (espace ouvert sec à couverture végétale spéciale) se subdivise en 4 sous-catégories (Landes, Toundra sèche, Pâturages et Autres). Ce découpage ne permet pas de classer de façon pertinente les paysages méditerranéens, rangés dans la sous-catégorie Autres. Pour inclure certains types de paysages (savane, chapparal, prairies), il est nécessaire de diviser la catégorie Autres sur un quatrième niveau (d'où un déséquilibre entre les catégories à ce niveau) ou de restructurer toute la catégorie.

Analyse discriminante

L'analyse discriminante, une autre approche envisageable, cherche à optimiser la séparation des classes définies a priori et à minimiser la variance au sein des classes grâce à l'utilisation de variables descriptives. Elle permet d'affecter des objets individuels à des classes mais requiert des données numériques.

3.3.2. Les systèmes non hiérarchiques

Noyaux élémentaires et intersection systématique

Cette méthode empirique vise à extraire des différents systèmes de classification existants les éléments de base qui établiront les catégories principales. Elle implique le regroupement des nomenclatures en vigueur dans les domaines considérés, leur confrontation et leur analyse afin de répondre aux besoins des utilisateurs. Les nomenclatures sont comparées en partant du principe que les agrégats importants peuvent être communs à plusieurs utilisateurs ou approches. Les agrégats généraux, acceptés normalement dans diverses nomenclatures, constituent la base (noyau) et les articles (qui peuvent être associés à plusieurs agrégats selon les nomenclatures) sont les éléments périphériques (limites). À titre d'exemple, les pâtures temporaires et artificielles seraient classées dans la catégorie « terres arables » dans les nomenclatures agricoles, mais appartiendraient à d'autres catégories selon d'autres utilisateurs (POIRET 1997). Le nombre de nomenclatures prises en compte doit être limité ; en effet, le nombre de noyaux identifiés est inversement proportionnel à celui des classifications, puisque la probabilité de découvrir des objets n'appartenant pas au noyau augmente avec la multiplication des perspectives.

L'intersection systématique des nomenclatures constitue une approche similaire. RADERMACHER (1988) propose une intersection systématique de deux nomenclatures de base relatives respectivement à l'occupation et à l'utilisation des sols. Le résultat est une table dans laquelle 9 occupations et 14 utilisations des sols ont fait l'objet d'un croisement systématique, générant 126 classes dont 75 étaient vides (des classes sans fondement telle que « eau utilisée à des fins sylvicoles »). Sur les 51 classes restantes, 9 groupes ont été retenus à des fins spécifiques (principalement écologiques), parmi lesquels 3 classes avaient purement trait à l'occupation des sols (aspect jugé fondamental pour les surfaces boisées, d'autres zones naturelles et les surfaces humides, indépendamment de leur utilisation). Les six autres associent occupation et utilisation, ou renvoient à une autre dimension ; par exemple, la catégorie « Zones sans objet écologique ».

Approches systématiques avec classificateurs

Plutôt qu'un système universel de classification pour l'occupation et l'utilisation des sols, l'important est de développer des outils simplifiant le lien entre les systèmes ; ces outils devraient être autant que possible affranchis des diverses contraintes liées à l'établissement des nomenclatures et à ce domaine spécifique. Ce pourrait être un système combiné, utilisé comme base commune : un simple ensemble de caractéristiques nécessaires à la description des objets. Celles-ci, une fois identifiées et spécifiées de manière uniforme, pourraient être agencées pour définir les objets et les groupes d'objets dans tous les systèmes.

Décrire un objet consiste à prendre en compte ses caractères (EKHOLM 1996), lesquels peuvent avoir plusieurs expressions (états). Il existe différents types de caractères : les caractères à deux états (présent / absent, positif / négatif, +/-, A/non A, 1/0), les caractères multi-états (états quantitatifs : continu ou discret) et les états qualitatifs (terminologiques). Ils peuvent être utilisés dans la règle de classification des objets à l'intérieur d'un système spécifique. Le tableau 3.1, qui prend l'exemple des éléphants, des baleines et des fourmiliers, illustre cette chaîne décisionnelle (DUHAMEL 1998).

	ANIMALITÉ	TAILLE	COULEUR DE LA PEAU	PRÉSENCE DE TROMPE	NOM (IMAGE) en français	CODE NOMENCLATURE X
	OUI	GRANDE	GRIS	OUI	ÉLÉPHANT	AH B200
	OUI	GRANDE	GRIS	NON	BALEINE	AV Z300
	OUI	MOYENNE	NOIR et BLANC	OUI	FOURMILIER	AB K541

Objet	Concept (standardisé par un ensemble de caractères)	Image

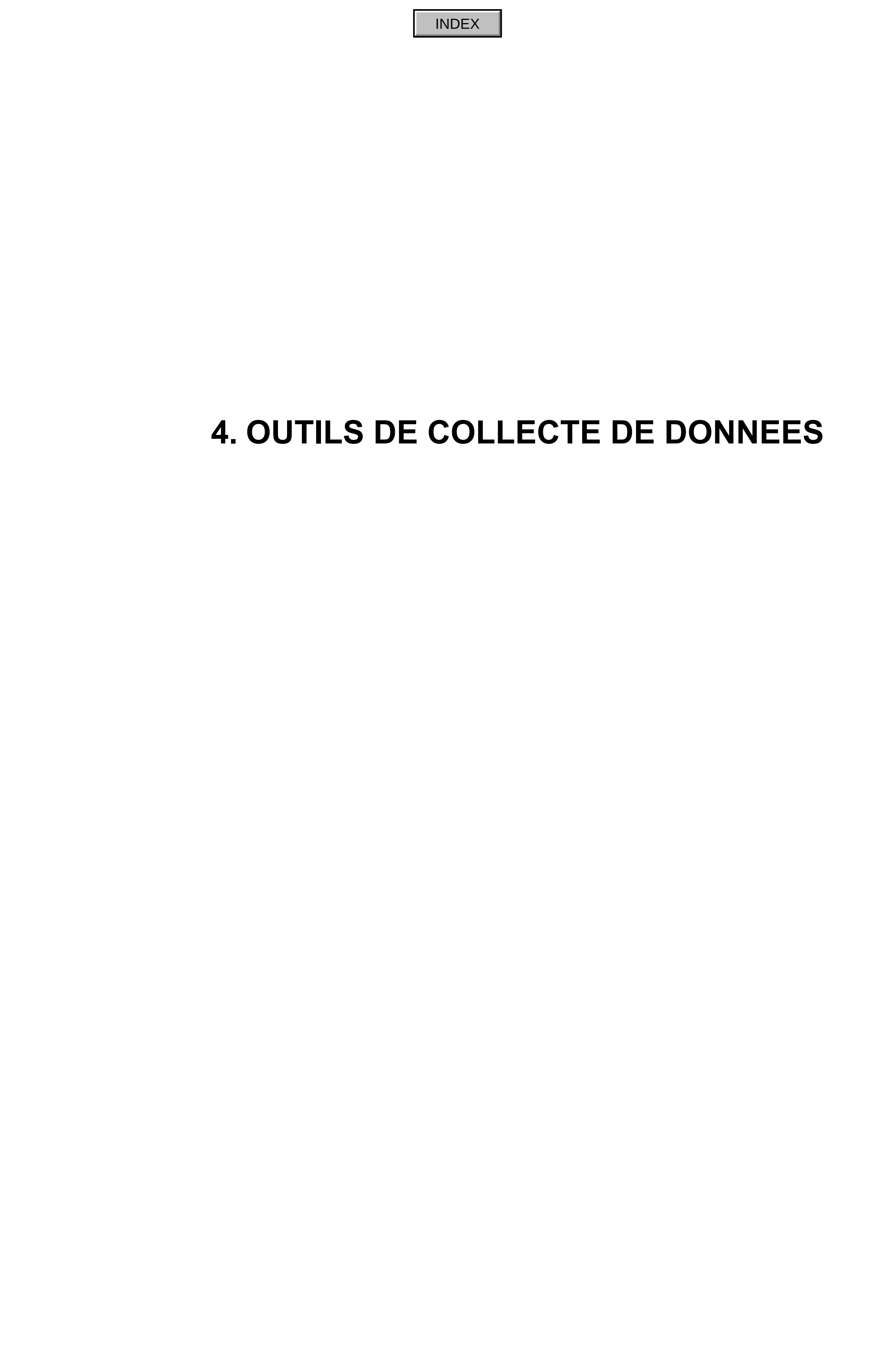
Tableau 3.1 : Exemple d'application de classificateurs

La sélection arbitraire de quatre propriétés illustre le principe de la description des objets par des caractères pertinents : un caractère à deux états, ANIMALITÉ (OUI ou NON) ; un caractère à trois états, TAILLE (GRANDE, MOYENNE, PETITE) ; un caractère multi-état, COULEUR DE LA PEAU, et un caractère à deux états, PRÉSENCE DE TROMPE (OUI ou NON). La réunion des quatre propriétés suivantes [ANIMALITÉ : OUI ; TAILLE : GRANDE ; COULEUR DE LA PEAU : GRISE ; PRÉSENCE DE TROMPE : OUI] définit les objets appelés « ÉLÉPHANTS ». Ces caractères peuvent être utilisés systématiquement comme **critères de diagnostic ou classificateurs**. Les classificateurs sont des représentations conceptuelles des caractères des objets, c'est-à-dire une règle de décision qui spécifie la classe attribuée à chaque observation (objet). Les classificateurs représentent des propriétés caractéristiques répondant aux objectifs de la nomenclature.

Cette approche peut servir à établir des systèmes de classification (comme c'est le cas, par exemple, en taxinomie des sols). La question de la hiérarchie des classificateurs est fondamentale ; elle signifie qu'il est possible de définir des règles *a priori* pour séparer les catégories en sous-catégories, et ainsi de suite. Cette méthode *a priori* génère un schéma hiérarchique proposant des pondérations inégales. Là encore, nous sommes confrontés à une approche traditionnelle des classifications hiérarchiques où des découpages successifs, fondés sur des caractéristiques/caractères « objectifs », suivent ce schéma : une structure dans laquelle un rang est attribué à chaque niveau de découpage correspondant à une série de caractères. La solution consiste à remplacer les schémas hiérarchiques rigides par un système combiné uniforme reposant sur une base commune de classificateurs de poids égal ; cette base se contente de regrouper les caractéristiques nécessaires à la description des objets. Ces caractéristiques, une fois identifiées et spécifiées de manière uniforme, permettent de définir les objets et les groupes d'objets de toutes les nomenclatures possibles. Un tel système peut difficilement être développé sans proposer de nombreux classificateurs pour opérer une distinction entre tous les

objets relatifs à l'occupation et à l'utilisation des sols. Une hiérarchisation est généralement introduite pour limiter le nombre de classificateurs. Deux approches successives sont proposées :

- des classificateurs généraux de rang supérieur qui peuvent être combinés au gré de l'utilisateur (classification à facettes) ;
- des descripteurs de rang inférieur qui sont spécifiques à un domaine particulier (adoption d'un principe d'économie pour supprimer les redondances et limiter le nombre des classificateurs).



4. OUTILS DE COLLECTE DE DONNEES

4. OUTILS DE COLLECTE DE DONNEES

Ce chapitre présente les différents outils de collecte permettant de recueillir les premières informations sur l'occupation et l'utilisation des sols. Il décrit les principes fondamentaux, les concepts sous-jacents et les différentes approches conceptuelles, en précisant brièvement leurs avantages et leurs limites. Ces explications techniques sont indispensables à l'évaluation des systèmes d'information et à la qualité du retour d'information.

4.1. Imagerie satellitaire obtenue par télédétection

4.1.1. Définition

« La télédétection est la technique qui, par l'acquisition d'images, permet d'obtenir de l'information sur la surface de la Terre sans contact direct avec celle-ci. La télédétection englobe tout le processus qui consiste à capter et à enregistrer l'énergie d'un rayonnement électromagnétique émis ou réfléchi, à traiter et à analyser l'information, pour ensuite mettre en application cette information ». (Centre canadien de télédétection (CCT) : Notions fondamentales de la télédétection 1999).

4.1.2. Principes

Le principe de la télédétection peut se résumer comme suit :

La surface de la Terre est illuminée par un large spectre de radiations électromagnétiques provenant du soleil. Le tableau 4.1 présente les différentes composantes du spectre électromagnétique, qui va de l'ultraviolet à l'infrarouge.

Spectre		**Longueur d'onde en µm (~10^{-6} m)**
Ultraviolet :		**< 0,3**
Partie visible du spectre	**Violet :**	**0,4 – 0,446**
	Bleu :	**0,446 – 0,500**
	Vert :	**0,500 – 0,578**
	Jaune :	**0,578 – 0,592**
	Orange :	**0,592 – 0,620**
	Rouge :	**0,620 – 0,7**
IR réfléchi		**0,7 – 3**
IR thermique		**3 - 100**

Tableau 4.1 : Le spectre électromagnétique

Le rayonnement solaire interagit avec l'atmosphère et n'atteint pas en totalité la surface de la Terre. Une grande partie de l'énergie est absorbée. L'ozone stratosphérique, par exemple, absorbe la majeure partie du rayonnement ultraviolet. L'eau contenue dans l'atmosphère est responsable de l'absorption de certaines composantes du rayonnement infrarouge. À l'exception de certaines longueurs d'onde, l'atmosphère est surtout transparente, notamment dans la région visible du spectre électromagnétique. Ce sont ces « fenêtres atmosphériques », par lesquelles la transmission de l'énergie n'est effectivement pas perturbée, qui sont utilisées en télédétection.

Tous les objets sur la surface terrestre (cibles) interagissent avec le rayonnement. Les cibles réfléchissent, transmettent ou absorbent les ondes électromagnétiques qui leur parviennent. Le processus dépend de la structure physique et chimique de la cible et de la longueur d'onde impliquée.

La portion réfléchie du spectre est la plus importante pour les applications de télédétection en rapport avec le sol. Selon les longueurs d'onde, les cibles réfléchissent le rayonnement de manière spécifique, et parfois unique (voir la figure 4.1).

Cette signature spectrale caractéristique des objets permet leur identification par la télédétection. En comparant les signatures de différents objets à la surface de la Terre, nous pouvons les distinguer les uns des autres.

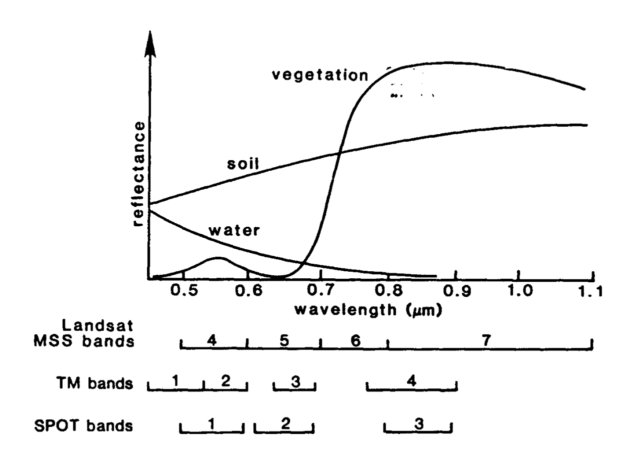


Figure 4.1 : Signature spectrale de quelques surfaces courantes et bandes spectrales des capteurs satellites

Examinons la signature spectrale de quelques cibles de la surface de la Terre :

- **La végétation** : la chlorophylle, une molécule que nous retrouvons à l'intérieur des feuilles, absorbe fortement le rayonnement aux longueurs d'onde du rouge et du bleu, mais réfléchit le vert. Les feuilles, qui contiennent un maximum de chlorophylle en été, sont donc plus vertes pendant cette saison. En automne, les feuilles qui contiennent alors moins de chlorophylle, absorbent moins de rouge, et paraissent donc rouges ou jaunes. La structure interne des feuilles saines agit comme un excellent réflecteur diffus pour les longueurs d'onde du proche infrarouge. Le ratio proche infrarouge/rouge révèle de nombreuses informations sur la végétation et il est utilisé à des fins de surveillance dans ce domaine. Il est possible d'identifier les plantes grâce à leurs propriétés de réflexion.
- **L'eau** : l'eau absorbe davantage les grandes longueurs d'onde du rayonnement visible et du proche infrarouge. Ainsi, l'eau paraît généralement bleue ou bleu-vert car elle réfléchit davantage les petites longueurs d'onde, elle paraît encore plus foncée si elle est observée sous les longueurs d'onde du rouge ou du proche infrarouge. Lorsque les couches supérieures de l'eau contiennent des sédiments en suspension, la transmission diminue, la réflexion augmente et l'eau paraît plus brillante.
- **Les sols et les minéraux** : les propriétés de réflexion des sols présentent des caractéristiques spectrales plus marquées. La réflexion dépend surtout de la composition minérale, de la granulométrie, de l'eau et du contenu organique du sol. Plus le sol est sec et pur, et plus le rayonnement est faible dans cette région du spectre.

Les signatures spectrales peuvent être très variables pour la même sorte de cible et peuvent aussi varier dans le temps et dans l'espace. Par ailleurs, elles sont influencées par les conditions atmosphériques puisque la transmission de l'énergie radiative dépend fortement de la vapeur d'eau et de la poussière contenues dans l'atmosphère.

Pour interpréter correctement l'interaction du rayonnement électromagnétique avec la surface, il est important de savoir où regarder dans le spectre et de comprendre les facteurs qui influencent la signature spectrale de la cible.

4.1.3 Technique de mesure : le balayage

Les capteurs enregistrent les données grâce à un système à balayage. Celui-ci utilise un capteur dont le champ de vision instantanée (CVI) est étroit, mais qui balaie la surface de façon à en produire une image bidimensionnelle (image matricielle, figure 4.2).

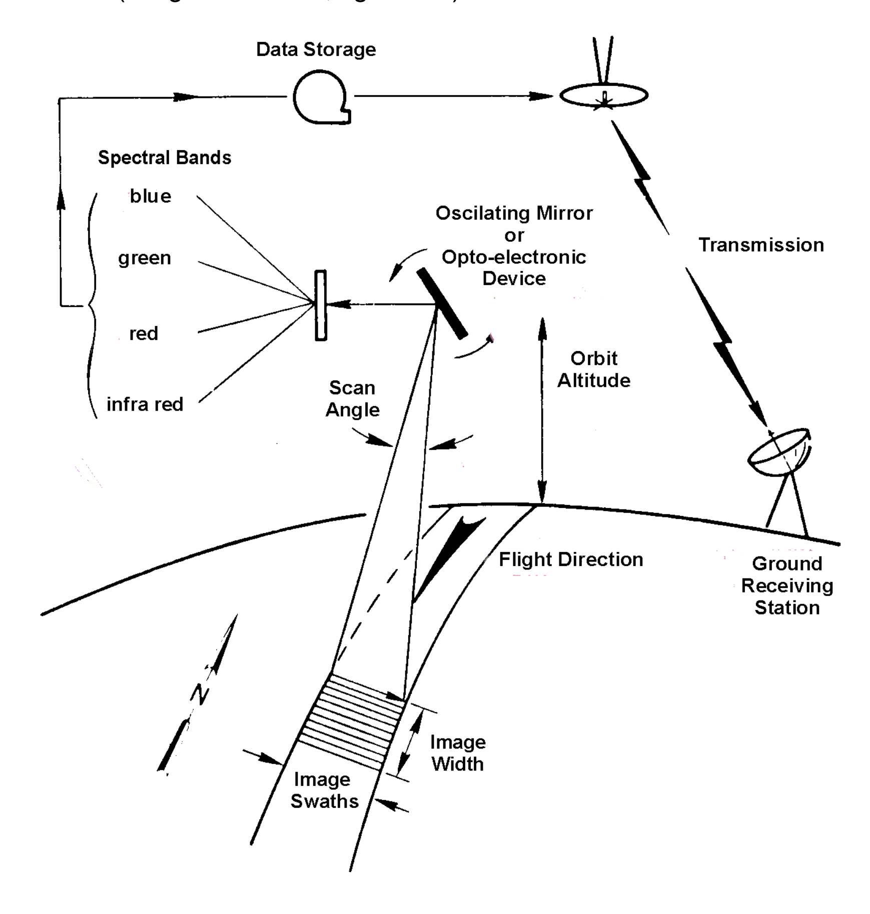


Figure 4.2 : Principe de l'enregistrement de données – le balayage

Les systèmes à balayage peuvent être utilisés sur des plates-formes aériennes ou spatiales, et opèrent essentiellement de la même façon. Un système à balayage qui utilise plusieurs longueurs d'onde (bandes spectrales) pour mesurer simultanément l'énergie réfléchie (ou émise) est appelé un balayeur multispectral (BMS). L'espacement des cellules tramées (éléments d'image ou pixels) et la région des longueurs d'onde dépendent des spécifications techniques du capteur (résolutions spatiale et spectrale).

Un système à balayage possède des avantages marqués sur les systèmes photographiques. Ces derniers sont limités à l'utilisation de bandes spectrales dans le visible et dans le proche infrarouge, tandis que les BMS peuvent aussi utiliser l'infrarouge thermique. Les BMS ont aussi une résolution spectrale supérieure aux systèmes photographiques. Les systèmes photographiques multispectraux utilisent des lentilles distinctes pour détecter simultanément chacune des bandes spectrales. Les systèmes photographiques enregistrent l'énergie détectée au moyen de processus photochimiques pour lesquels il est plus difficile de faire des mesures et de contrôler la fiabilité. Comme les données provenant des BMS sont enregistrées électroniquement, il est plus facile de déterminer la quantité d'énergie détectée et d'augmenter la résolution radiométrique.

Plates-formes satellites

Voici une courte description des satellites et capteurs d'observation les plus connus :

- Landsat

Lancé en 1972, le satellite d'observation américain Landsat fournit des images multispectrales de la surface de la Terre. Les trois premiers satellites, Landsat-1 à Landsat-3, opérationnels jusqu'en 1983, étaient équipés du système MSS (*Multi Spectral Scanner*) : 4 bandes, avec une fauchée de 185 km x 185 km et une résolution spatiale de 80 m. Landsat 4 (lancé en 1982) et Landsat 5 (en 1984) portent en plus le capteur TM (*Thematic Mapper*), qui inclut 7 bandes spectrales avec une fauchée de 185 km x 185 km et une résolution spatiale de 30 m. Quant à Landsat-7, opérationnel depuis avril 1999, il comprend un capteur TM amélioré (ETM) qui, outre les bandes multispectrales similaires à celles de Landsat-5, peut capter en mode panchromatique avec une résolution spatiale de 15 m. Les volumineuses archives d'images récoltées pendant toute la phase opérationnelle du programme satellite Landsat offrent une vue rétrospective et permettent d'analyser les changements.

- SPOT

Le programme français SPOT (Système Probatoire d'Observation de la Terre) a démarré en 1986 avec SPOT 1, suivi par SPOT 2 en 1990 et par SPOT 3 en 1993. Tous les satellites sont en orbite héliosynchrone polaire à une altitude de 830 km et peuvent capter en mode multibande (3 bandes spectrales) HRV (Haute Résolution Visible) qui offre une résolution spatiale de 20 m, ainsi qu'en mode panchromatique (une seule bande) avec une résolution spatiale de 10 m. SPOT 4, lancé en 1998, porte également l'instrument « Végétation », pour des prises de vue sur 4 bandes (bleu, rouge, proche infrarouge et infrarouge court) avec une résolution spatiale de 1 km et une largeur de 2 250 km. Le lancement de SPOT 5 est prévu en l'an 2000 ; il emportera un capteur générant des images panchromatiques avec une résolution spatiale de 5 m.

- IRS

Le premier satellite de la série IRS (*Indian Remote Sensing*), IRS-1A, a été lancé en 1988. La dernière plate-forme, IRS-1D, a trois capteurs : une caméra de haute résolution panchromatique avec une résolution spatiale de 5,8 m, le capteur à quatre bandes LISS (*Linear Imaging Self-scanning Sensor*) avec une résolution de 23,5 m, et le capteur à deux bandes WiFS (*Wide Field of View*) avec une résolution de 188 m.

- IKONOS

Premier satellite commercial à haute résolution, IKONOS est opérationnel depuis fin 1999. Depuis son orbite située à 681 km d'altitude, ses deux capteurs indépendants balaient la surface de la Terre sur une bande de 11 km de large pouvant atteindre 1 000 km de long. Avec une résolution spatiale de 1 m (mode panchromatique) et de 4 m (mode multispectral), il fournit des images de haute qualité, proche de celle des photographies aériennes, et ouvre de nouvelles perspectives d'application de l'imagerie par télédétection dans de nombreux domaines.

- NOAA - AVHRR

La NOAA (*National Oceanic and Atmospheric Administration*) opère aussi une série de satellites équipés du capteur AVHRR (*Advanced Very High Resolution Radiometer*), conçus à l'origine pour être utilisés en météorologie. Leurs spécifications techniques sont donc assez différentes de celles des satellites d'observation de la Terre (Landsat, SPOT ou IRS). Le capteur AVHRR capte le rayonnement électromagnétique du visible et du proche IR, ainsi que l'énergie émise (température) dans la région du moyen et du long IR. Les informations sur la température, notamment sur les nuages, présentent un intérêt spécifique pour les météorologues. Avec sa résolution spatiale de 1,1 km et son grand angle de prise de vue (55°), le capteur AVHRR a une fauchée d'environ 2 000 km (Landsat TM : 185 km). Les propriétés techniques, et le fait qu'au moins deux satellites soient opérationnels en même temps, garantissent une couverture quotidienne et une surveillance continue du globe terrestre, qui s'avèrent essentielles pour les prévisions et autres domaines météorologiques.



Figure 4.3 : Image IRS III de la vallée du Rhin au nord de Karlsruhe, le 2 novembre 1997 (Proche IR, Rouge, vert = RGB ; résolution spatiale : 23 m) ; Source : Euromap, Neustrelitz (www.euromap.de)



Figure 4.4 : Image panchromatique IRS-1C de l'aéroport Schiphol d'Amsterdam, le 17 mai 1998 (résolution spatiale : 5,8 m) Source : Euromap, Neustrelitz (www.euromap.de)

- RADARSAT, ERS

Du point de vue de la conception et de la technique, les systèmes satellites radar sont très différents des plates-formes présentées jusqu’ici. Contrairement aux capteurs, qui mesurent le rayonnement solaire réfléchi ou émis par la surface terrestre (systèmes passifs), le radar est un système actif qui fournit sa propre source d’énergie électromagnétique et mesure la portion d’énergie renvoyée par la cible. Les propriétés structurelles des cibles déterminent les caractéristiques de la réflexion et permettent leur identification. L’avantage du radar est sa capacité de pénétration à travers les nuages et de détection dans presque toutes les conditions atmosphériques. Il est peu utilisé par les applications cartographiques relatives à l’occupation et à l’utilisation des sols (c’est la raison pour laquelle ce système n’est pas traité plus en détail ici). Il est en revanche plus fréquemment utilisé en géologie, pour la détection des glaces de mer ou des marées noires. Le satellite canadien RADARSAT et les satellites européens ERS-1 et ERS-2 sont deux systèmes radar opérationnels.

4.1.4 Images de télédétection : propriétés

Les possibilités d’application des images obtenues par télédétection sont surtout déterminées par les spécifications techniques des capteurs et par l’orbite des satellites. Nous développons dans ce chapitre les points les plus importants à ce sujet.

Résolution spectrale

Comme indiqué plus haut, une cible à la surface de la Terre se caractérise par sa réflectance spectrale à différentes longueurs d’onde. Les capteurs mesurent le rayonnement réfléchi à plusieurs intervalles, appelés bandes spectrales ou canaux spectraux, afin de percevoir ces différences. La capacité d’un capteur à identifier les cibles dépend en grande partie du nombre de bandes spectrales, c’est-à-dire de sa résolution spectrale.

La résolution spectrale décrit la capacité d’un capteur à utiliser de petites fenêtres de longueurs d’onde. Plusieurs instruments de télédétection peuvent enregistrer l’énergie reçue selon des intervalles de longueurs d’onde (bandes spectrales ou canaux spectraux) à différentes résolutions spectrales. Plus la résolution spectrale est fine, plus les fenêtres des différents canaux du capteur sont étroites, et plus la détection et la différenciation des objets sont précises. Des capteurs multispectraux plus développés, appelés capteurs hyperspectraux, sont capables de détecter des centaines de bandes spectrales très fines dans la portion du spectre des ondes électromagnétiques réunissant le visible, le proche infrarouge et l’infrarouge moyen. Leur grande résolution spectrale facilite la différenciation des caractéristiques d’une image basée sur la réponse différente dans chacune des bandes spectrales.

Le tableau 4.2 indique la résolution spectrale des principaux capteurs utilisés pour les applications d’observation de la Terre. Avec ses 7 bandes, Landsat TM (*Thematic Mapper*) assure la meilleure différenciation des objets ; SPOT et IRS n’utilisent respectivement que 3 et 4 bandes spectrales.

SPOT, IRS et le dernier capteur Landsat ETM fonctionnent également en mode panchromatique, balayant la surface terrestre avec un intervalle de longueurs d’onde assez grand (spectre visible) et une résolution spatiale supérieure à celle des bandes multispectrales.

Capteur	Résolution spectrale		Résolution spatiale	**Échelle maximale recommandée (approximation)**
	Canal	Domaine spectral (en microns)		
Landsat MSS	**MSS 1**	**0,5 – 0,6 (vert)**	**80 m * 80 m**	1/500 000e
	MSS 2	**0,6 – 0,7 (rouge)**		
	MSS 3	**0,7 – 0,8 (proche IR)**		
	MSS 4	**0,8 – 1,1 (proche IR)**		
Landsat TM/ ETM	**TM 1**	**0,45 – 0,52 (bleu)**	**30 m * 30 m**	1/200 000e
	TM 2	**0,52 – 0,60 (vert)**		
	TM 3	**0,63 – 0,69 (rouge)**		
	TM 4	**0,76 – 0,90 (proche IR)**		
	TM 5	**1,55 – 1,75 (IR court)**		
	TM 6	**10,4 – 12,5 (IR thermique)**	**120 m * 120 m**	
	TM 7	**2,08 – 2,35 (IR court)**		
Landsat ETM	**Panchromatique**	**0,52-0,9**	**15 m * 15 m**	**1/100 000e**
SPOT Pan	**Panchromatique**	**0,51 – 0,73 (bleu-vert-rouge)**	**10 m * 10 m**	**1/50 000e**
SPOT XS Multispectral	**Bande 1**	**0,50 – 0,59 (vert)**	**20 m * 20 m**	**1/100 000e**
	Bande 2	**0,61 – 0,68 (rouge)**		
	Bande 3	**0,79 – 0,89 (proche IR)**		
SPOT Végétation	**Canal 1**	**0,50-0,59 (vert)**	**1000 m * 1000 m**	**1/1 500 000e**
	Canal 2	**0,61-0,68 (rouge)**		
	Canal 3	**0,79-0,89 (proche IR)**		
	Canal 4	**1,58-1,75 (IR court)**		
IRS Pan	**Panchromatique**	**0,5 – 0,75 (bleu-vert-rouge)**	**5,8 m * 5,8 m**	**1/15 000e**
IRS LISS	**Bande 2**	**0,52 – 0,59 (vert)**	**23,5 m * 23,5 m**	**1/100 000e**
	Bande 3	**0,62 – 0,68 (rouge)**		
	Bande 4	**0,77 – 0,86 (proche IR)**		
	Bande 5	**1,55 – 1,7 (IR court)**		
IRS WiFS	**Bande 3**	**0,62 – 0,68 (rouge)**	**188 m * 188 m**	**1/500 000e**
	Bande 4	**0,77 – 0,86 (proche IR)**		
NOAA AVHRR	**Canal 1**	**0,58- 0,68 (rouge)**	**1100 m * 1100 m**	**1/1 500 000e**
	Canal 2	**0,725 – 1,1 (proche IR)**		
	Canal 3	**3,55 – 3,93 (IR moyen)**		
	Canal 4	**11,3 – 11,3 (IR thermique)**		
	Canal 5	**11,4 – 12,4 (IR thermique)**		
IKONOS	**Panchromatique**	**0,45 – 0,90**	**1 m * 1 m**	**1/5000e**
multispectral	**Canal 1**	**0,45 – 0,52 (bleu)**	**4 m * 4 m**	**1/15 000e**
	Canal 2	**0,52 – 0,60 (vert)**		
	Canal 3	**0,63 – 0,69 (rouge)**		
	Canal 4	**0,79 – 0,90 (proche IR)**		

Tableau 4.2 : Résolution spectrale, résolution spatiale et échelle recommandée (approximation) des principaux instruments de télédétection

Résolution spatiale

Les images de télédétection sont composées d'une matrice d'éléments appelés pixels. Le pixel est le plus petit élément de la mesure de la réponse spectrale. Sa taille varie en fonction des caractéristiques techniques des capteurs.

Les images sur lesquelles seuls les grands éléments sont visibles ont une résolution « grossière » ou « basse ». Les images à résolution fine ou élevée permettent l'identification d'éléments de plus petites dimensions. Les satellites commerciaux ont une résolution qui varie de quelques mètres à plusieurs kilomètres.



Figure 4.5 : Images satellites de Berlin (Allemagne) à différentes résolutions spatiales : en haut à gauche : MSS (80 m) ; en haut à droite : TM (30 m) ; en bas à gauche : SPOT multispectral (20 m) ; en bas à droite : SPOT panchromatique (10 m) (ALBERTS 1991)

La résolution spatiale a des implications importantes pour l'identification des objets en surface, l'échelle de l'analyse, la précision et la pertinence de la localisation. Plus la résolution est fine, plus les formes des objets sont précises et complètes, et plus il est possible de les identifier à partir de leur forme et de déterminer avec précision leur emplacement et leur étendue (voir la figure 4.5).

Pour pouvoir différencier un élément de la surface observée, l'élément en question doit être de dimension égale ou supérieure au pixel. S'il est plus petit, il ne sera généralement pas différencié puisque c'est l'énergie moyenne des éléments du pixel qui sera captée. Cependant, dans certaines conditions, un élément plus petit peut être détecté si sa réflexion domine celle des autres éléments présents dans le pixel.

La résolution spatiale des données doit être compatible avec les objectifs du projet. Les nouvelles avancées dans le domaine de la télédétection visent à obtenir des images à haute résolution spatiale, d'environ 1 à 2 mètres, permettant d'identifier de petits éléments.

Résolution radiométrique

La résolution radiométrique fait référence à la plage dynamique, c'est-à-dire à la gamme des longueurs d'onde à l'intérieur de laquelle un capteur est sensible. La gamme dynamique de la plupart des données satellitaires est de 7 bits ou 128 niveaux d'intensité (Landsat MSS, IRS), ou 256 niveaux (8 bits) pour Landsat TM, ETM.

Plus la résolution radiométrique d'un système de télédétection est fine, plus il est sensible à de petites différences dans l'intensité de l'énergie électromagnétique reçue. De nombreux outils logiciels sont conçus pour traiter des images de 8 bits ; pour les autres longueurs d'octets, un traitement spécial peut être nécessaire.

Résolution temporelle

En plus de la résolution spatiale et spectrale, l'autre concept important en télédétection est celui de la résolution temporelle. Celle-ci dépend de la période de passage au nadir d'un satellite, qui est le temps que prend un satellite pour effectuer un cycle orbital complet et pour observer de nouveau exactement la même scène à partir du même point dans l'espace (voir les figures 4.6 et 4.7).

Cette période est généralement de quelques jours (voir le tableau 4.3). La résolution temporelle absolue du système de télédétection est donc égale à cette période.

Plates-formes de télédétection	**Résolution temporelle**
IKONOS	4 jours
Landsat 1,2 et 3	18 jours (toutes les 251 révolutions)
Landsat 4,5 et 7	16 jours (toutes les 233 révolutions)
SPOT	26 jours
NOAA-AVHRR	1 jour

Tableau 4.3 : Résolution temporelle des principales plates-formes satellites

Toutefois, certaines régions de la surface peuvent être observées plus fréquemment puisqu'il y a chevauchement entre les couloirs-couverts adjacents et que ces zones de chevauchement deviennent de plus en plus grandes en s'approchant des pôles. Certaines plates-formes (SPOT, IKONOS, par exemple) ont aussi la possibilité de pointer leurs capteurs en direction du même point pour différents passages du satellite, séparés par des périodes de un à cinq jours. La résolution temporelle effective du satellite dépend donc d'une variété de facteurs, dont la grandeur de la zone de chevauchement entre les couloirs-couverts adjacents, la capacité du satellite et de ses capteurs, l'altitude de l'orbite et également la latitude des zones observées.

Le principal facteur influençant la résolution temporelle est le fait qu'une absence de couverture nuageuse est requise lors de la télédétection. Les chances de prendre des images d'une région donnée sans nébulosité sur des orbites successives sont limitées. Les chiffres du tableau 4.3 tiennent davantage de la théorie que de la pratique.

L'un des grands avantages de la télédétection satellitaire est sa capacité à amasser périodiquement de l'information d'une même région de la Terre. Les caractéristiques spectrales de la région observée peuvent changer avec le temps. La comparaison d'images multitemporelles permet de détecter ces changements. Par exemple, durant la période de croissance de la végétation, de nombreuses espèces se transforment continuellement et notre capacité à détecter ces changements subtils dépend de la fréquence avec laquelle les données sont recueillies.

En amassant des données périodiquement et de façon continue, il est possible de suivre les changements qui surviennent à la surface de la Terre, qu'ils soient naturels (comme le développement de la végétation ou l'évolution d'une inondation) ou de source humaine (comme le développement des milieux urbains ou la déforestation).

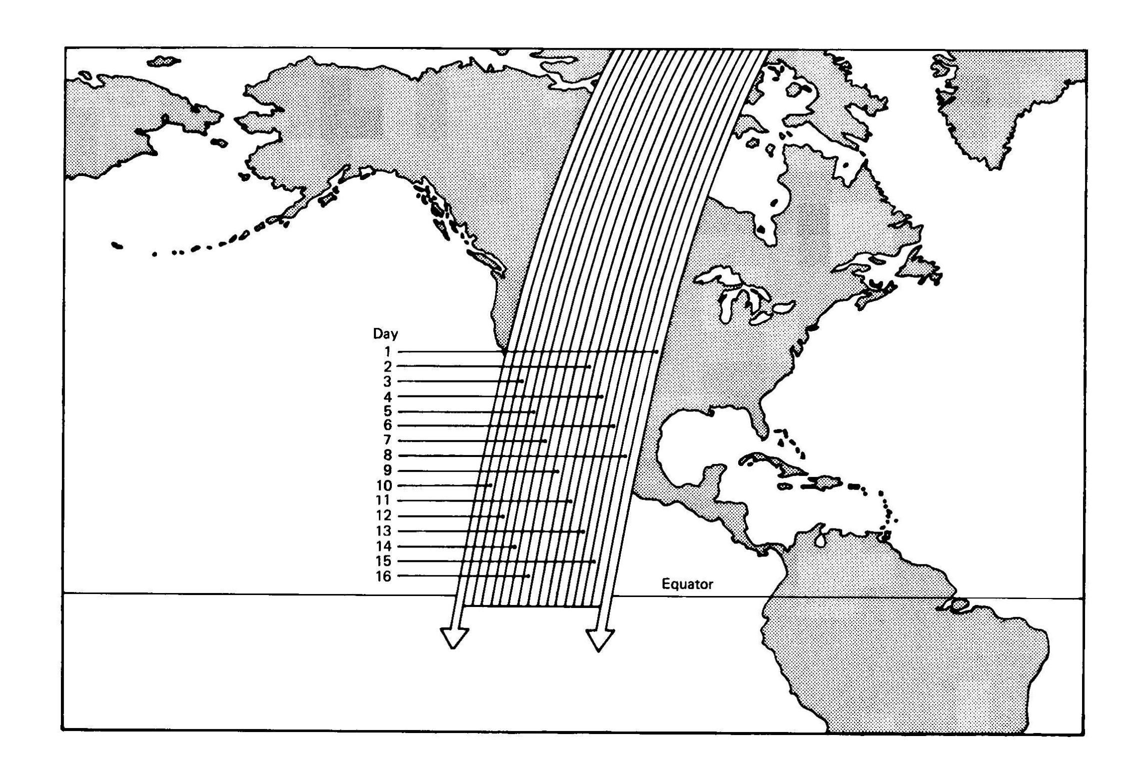


Figure 4.6 : Périodicité des couloirs-couverts adjacents de Landsat-4 et Landsat-5. Les fauchées adjacentes sont enregistrées à 7 jours d'intervalle (LILLESAND & KIEFER 1994)

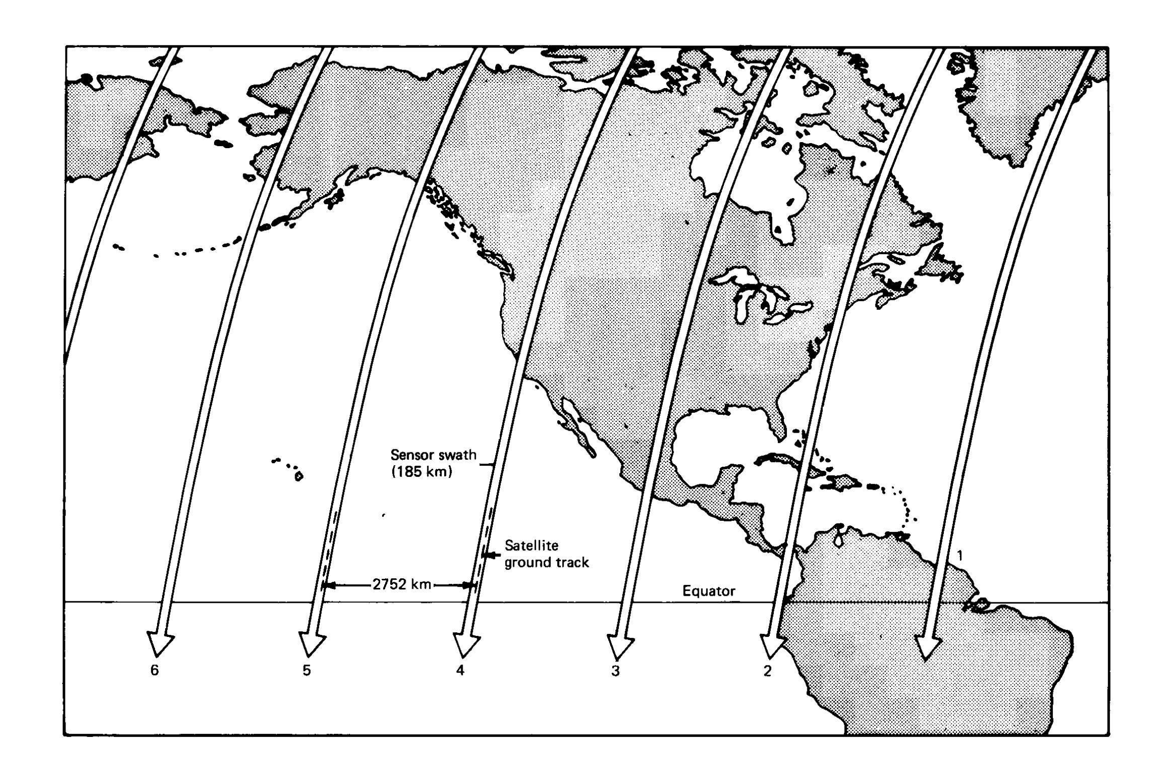


Figure 4.7 : Espacement entre des couloirs-couverts adjacents de Landsat à l'équateur (LILLESAND & KIEFER 1994)

4.1.5 Traitement numérique des images

La plupart des données de télédétection étant enregistrées au format numérique, presque toutes les interprétations et analyses d'images requièrent une partie de traitement numérique. Rares sont les données directement exploitables. Plusieurs opérations de pré-traitement des images sont requises avant la phase d'analyse :

- Correction radiométrique
- Correction géométrique
- Rehaussement de l'image

Les corrections radiométriques sont nécessaires en raison des irrégularités du capteur et des perturbations atmosphériques. Un traitement est également requis pour pouvoir calculer des unités physiques (comme le rayonnement réfléchi ou émis). En particulier, la radiométrie doit être examinée avec soin quand des données de télédétection sont utilisées à des fins de surveillance.

Les corrections géométriques comprennent la transformation des données en vraies coordonnées (par exemple en latitude et longitude) sur la surface de la Terre.

Les fonctions de rehaussement visent à améliorer l'apparence de l'imagerie pour aider l'interprétation et l'analyse visuelles.

4.1.6 Analyse d'image

L'interprétation et l'analyse de l'imagerie de télédétection ont pour but d'identifier et de mesurer différentes cibles dans une image pour pouvoir en extraire l'information utile. En télédétection, une cible est définie comme étant toute structure ou objet observable dans une image. La cible doit être distinctive aux niveaux spatial et spectral, c'est-à-dire qu'elle doit contraster avec les structures avoisinantes. L'analyse de l'image est visuelle (analyse d'image analogique, telle qu'une interprétation visuelle) et/ou numérique.

L'interprétation et l'analyse visuelles datent du tout début de la télédétection avec l'interprétation de photos aériennes. Le traitement et l'analyse numériques sont plus récents avec la venue des enregistrements numériques des données de télédétection et le développement des ordinateurs. L'analyse numérique est réalisée pour prendre en compte l'intégralité des données spectrales (bases de données multicanaux). Différentes méthodes et approches ont été développées à des fins spécifiques.

La classification numérique des images est appliquée dans le domaine de l'occupation des sols. À partir de la signature spectrale, l'objectif est d'assigner une classe particulière ou catégorie (forêts de conifères, prairies) à chacun des pixels d'une image. Le résultat de cette classification est une représentation thématique.

L'analyse ne repose pas uniquement sur le contenu de l'image numérique mais aussi sur des informations connexes comme des cartes topographiques ou thématiques, et des modèles numériques du relief. Il ne faut surtout pas oublier que les études de réalité de terrain sont indispensables à la validation des résultats. L'utilisation des images satellitaires ne dispensent pas des enquêtes de terrain.

Les techniques analogiques et numériques pour l'interprétation des données de télédétection ont leurs avantages et inconvénients respectifs. En général, l'interprétation visuelle, basée sur des documents papier, requiert peu ou pas d'équipement spécialisé, tandis que le traitement et l'analyse numériques nécessitent au moins un PC standard et un logiciel spécifique (Erdas Imagine, ER Mapper, etc.). L'interprétation visuelle se limite souvent à un seul canal de données ou à une seule image à la fois, compte tenu de la difficulté d'effectuer une interprétation visuelle avec plusieurs images. Les ordinateurs étant en mesure de traiter des images plus complexes, l'analyse numérique peut exploiter des données provenant de plusieurs canaux.

Dans cette optique, l'analyse numérique est utile à l'analyse simultanée de plusieurs bandes spectrales et elle peut traiter des banques de données plus rapidement qu'un interprète humain. Contrairement à l'analyse informatique, le facteur humain permet l'intégration de plusieurs informations contextuelles reposant sur les connaissances a priori du domaine d'étude. L'interprétation humaine est un procédé subjectif, ce qui veut dire que les résultats peuvent varier d'une personne à l'autre. L'analyse numérique, puisqu'elle est basée sur la manipulation de nombres par un ordinateur, est plus objective et fournit des résultats reproductibles. Cependant, la précision des résultats provenant de traitements numériques peut être difficile, puisqu'elle doit être validée par les analystes.

Rappelons que les analyses visuelles et numériques de l'imagerie de télédétection ne s'excluent pas mutuellement. Les deux méthodes ont chacune leurs mérites, et c'est en combinant les deux que se font la plupart des analyses. La décision finale quant à l'importance et à l'utilité de l'information extraite à la fin du processus d'analyse, revient toujours aux analystes.

4.1.7 Avantages et inconvénients de l'imagerie de télédétection

Il ne fait aucun doute que la télédétection contribue à une meilleure compréhension des processus à la surface de la Terre.

Les données de télédétection permettent d'avoir une vue synoptique de vastes superficies. La position, la répartition et les relations spatiales des éléments de la surface terrestre sont mises en évidence. Les relations spatiales peuvent ainsi être analysées.

Les capteurs couvrent une plus large portion du spectre que l'œil humain, permettant ainsi la détection et l'identification des caractéristiques environnementales de la surface terrestre ou de l'atmosphère, en particulier quand ils portent sur une largeur de bande très spécifique.

Les prises de vue répétées de la même zone constituent une source d'information unique pour les activités de suivi et de détection des changements. L'utilisation des données de télédétection à des fins de surveillance s'accompagne d'un travail long et coûteux car il nécessite l'adoption d'une méthodologie et des fonctions de pré-traitement pour apporter les corrections géométriques et radiométriques requises.

L'imagerie satellitaire joue un rôle important dans l'aménagement du territoire, notamment lorsqu'il n'existe aucune donnée de base, et l'échelle relativement petite s'avère suffisante.

L'utilisation des données satellitaires et la capacité de détection et d'identification (des classes d'occupation des sols, par exemple) dépendent des résolutions spectrale et spatiale des capteurs. La résolution spatiale détermine l'échelle. La télédétection permet en principe d'établir des cartes au 1/50 000e ou 1/100 000e. Dans un paysage très urbanisé, une résolution spatiale de 20 m, par exemple, ne permet pas une différenciation suffisante des objets dans la zone considérée. Du fait de sa résolution spatiale relativement grossière, l'imagerie satellitaire produit des cartes dont les échelles ne sont pas toujours appropriées. Les nouveaux systèmes à haute résolution, comme IKONOS, offrent une résolution nettement plus fine et permettent d'établir des cartes au 1/5 000e.

4.2. Photographies aériennes

Les photographies aériennes sont des images en noir et blanc ou en couleur de la surface terrestre, prises par un appareil photo embarqué sur un avion, un hélicoptère ou un ballon (NASA, 1999).

En télédétection, il est important de faire la distinction entre « image » et « photographie ». La première est une représentation graphique, quels que soient la longueur d'onde ou le dispositif de télédétection qui ont été utilisés pour capter et enregistrer l'énergie électromagnétique. La seconde désigne spécifiquement toute image captée et enregistrée sur une pellicule photographique (Centre canadien de télédétection, 1999).

4.2.1. Principes

Contrairement aux capteurs de télédétection, un appareil photographique enregistre l'énergie électromagnétique réfléchie ou émise par la surface de la Terre. Le processus photographique utilise une réaction chimique sur une surface sensible à la lumière pour capter et enregistrer les variations d'énergie. Les photographies enregistrent habituellement les longueurs d'onde entre 0,3 et 0,9 microns (les portions visible et infrarouge réfléchi). En général, l'appareil est placé à la verticale du sol et orienté vers le centre de la photo (nadir).

La taille de la photographie, c'est-à-dire la surface enregistrée, dépend des caractéristiques techniques et optiques de l'appareil (longueur focale, taille de champ) et de l'altitude de la plate-forme.

L'échelle des photographies aériennes va du 1/60 000e (petite) au 1/1 000e (grande).

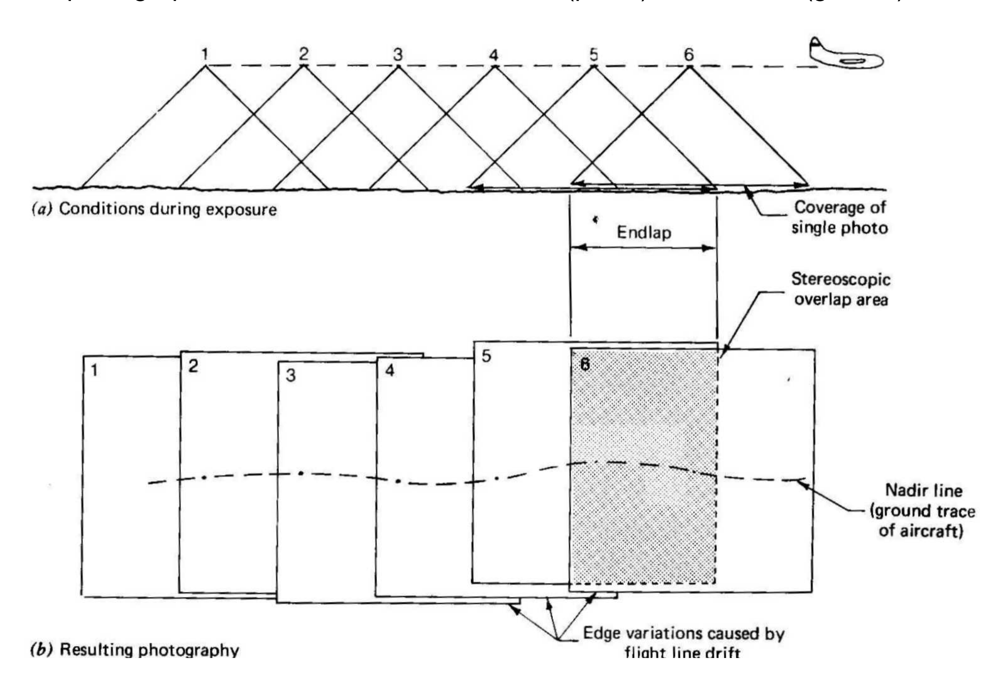


Figure 4.8 : Couverture sur une bande photographique (LILLESAND & KIEFER 1994)

Une photographie peut être présentée et affichée au format numérique, par simple balayage, en divisant l'image en petits morceaux de taille et de forme égales, que nous nommons pixels. La luminosité de chaque pixel est représentée par une valeur numérique.

Bien que les cartes comme les photographies aériennes correspondent à une vue « d'en haut », elles ne doivent pas être confondues. Les premières sont des représentations orthogonales de la surface terrestre, c'est-à-dire qu'elles sont exactes sur le plan directionnel et géométrique. Les secondes, en revanche, présentent une forte distorsion radiale. La topographie est déformée et, tant que les distorsions n'ont pas été corrigées, les mesures faites à partir d'une prise de vue ne sont pas fiables.

En principe, des bandes de terrain sont survolées avec un recouvrement longitudinal (entre clichés successifs) de 60 % et un recouvrement latéral (entre deux images de deux bandes adjacentes) variant de 20 à 40 % (voir la figure 4.8). En raison du recouvrement latéral, les éléments présents sur la surface sont photographiés sous deux angles différents. Cette parallaxe permet d'obtenir une vue tridimensionnelle (stéréoscopique) du terrain ainsi que des informations sur la hauteur des objets ou l'altimétrie du relief (lignes de contour).

4.2.2. Types de photographie aérienne

Noir et blanc

La pellicule des photographies monochromatiques diffère légèrement d'un film ordinaire.

La première pellicule noire et blanc fiable a été mise au point au début du siècle. L'essor de l'aéronautique a marqué le début de l'ère des photographies aériennes. Grâce à cette longue tradition, les agences nationales de cartographie possèdent des archives importantes de clichés aériens qui révèlent des détails sur les modifications du paysage.



Figure 4.9 : Exemple d'une série de clichés aériens en noir et blanc de Westkreuz, Allemagne 1952 – 1959 – 1964 (HANSA LUFTBILD www.hansaluftbild.de)

Infrarouge couleurs (IRC)

L'infrarouge couleurs est souvent appelé film « en fausses couleurs ». Contrairement au film noir et blanc et au film couleur ordinaire, l'émulsion est sensible au rayonnement infrarouge. Une végétation verte réfléchit très fortement le rayonnement infrarouge et apparaît en rouge vif sur les clichés IRC. Les objets dont la couleur naturelle est le rouge ressortent en vert et les objets verts (à l'exception de la végétation) en bleu. La photographie IRC est essentiellement utilisée dans les études de la végétation et en sylviculture (voir la figure 4.10).

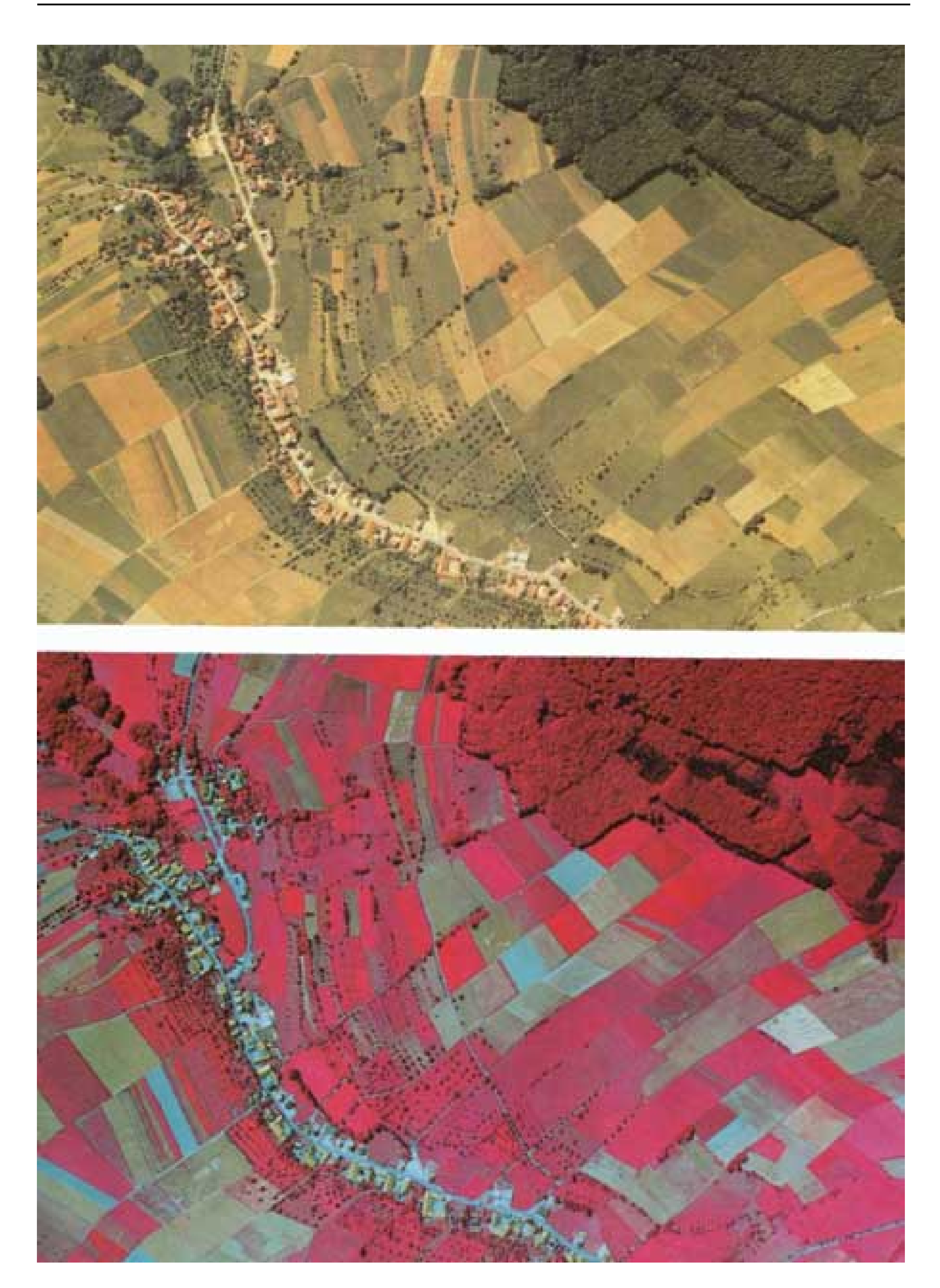


Figure 4.10 : Exemple de photographies aériennes couleur et IRC d'Odenwald, Allemagne, le 30 juillet 1967 ; l'échelle d'origine est au 1/14 000e (ALBERTZ 1991)

Orthophotographie

Une orthophoto (monochromatique, couleur ou IRC) est une prise de vue aérienne dont la géométrie a été corrigée afin de supprimer les distorsions causées par l'angle de l'appareil et le relief du terrain. Les orthophotos sont mises à une échelle précise et, grâce au géoréférencement, elles constituent une référence spatiale uniforme, dans un format harmonisé. Converties au format numérique, elles peuvent être intégrées à des systèmes d'information géographique (SIG) sans nécessiter d'autre traitement.

La richesse de leur contenu (espacement des pixels jusqu'à 0,25 cm) en fait une source d'information très précise.

Dans le cadre du SIGC (**S**ystème **I**ntégré de **G**estion et de **C**ontrôle), des archives importantes d'orthophotos numériques couvrant l'Italie, la Grèce, le Portugal, l'Irlande, le Danemark, la Belgique, la Finlande, les régions méridionales de la France et une superficie d'environ 400 000 km^2 en Espagne, ont été constituées. Les images monochromatiques balayées (au 1/40 000e) apportent aux agriculteurs une aide pour l'établissement de leurs formulaires de déclaration et sont utilisées par les autorités nationales à des fins de contrôle et de validation dans la lutte anti-fraude.

4.2.3. Principaux éléments d'interprétation des clichés aériens

Les photographies aériennes sont surtout interprétées visuellement. Les tons et les variations de teinte (gris ou couleur), ainsi que le patron qui en résulte, déterminent le contenu pictural d'un cliché aérien. Chacune de ces caractéristiques d'interprétation est décrite ci-dessous :

Le ton fait référence à la clarté relative ou couleur (teinte) des objets dans une image.

La forme fait référence au contour général des objets. Les formes géométriques régulières sont généralement les indices d'une présence et d'une utilisation humaines. Certains objets peuvent être identifiés uniquement à partir de leur forme.

La texture fait référence à l'arrangement et à la fréquence des variations de teintes dans des régions particulières d'une image (apparence rugueuse ou lisse). Cette impression est produite par des éléments trop petits pour être identifiés séparément. L'herbe, le ciment et l'eau renvoient une texture lisse, alors qu'une forêt peut résulter en une texture d'apparence rugueuse.

Le patron fait référence à l'agencement spatial des objets visiblement discernables. Les surfaces naturellement boisées et les vergers avec leurs arbres régulièrement disposés sont de bons exemples de patrons.

Le site fait référence à l'emplacement topographique ou géographique. Cette caractéristique est surtout importante pour l'identification des types de végétation et du relief.

L'association tient compte de la relation entre la cible d'intérêt et d'autres objets ou structures reconnaissables qui sont à proximité. Par exemple, une centrale nucléaire est (en principe) peu susceptible d'être implantée au beau milieu d'une zone d'habitation.

4.2.4. Avantages et inconvénients

Les photographies aériennes ne nécessitent aucun équipement matériel ou logiciel sophistiqué. En règle générale, la grande échelle de l'image permet d'identifier, de décrire et de délimiter parfaitement les objets les plus petits. En raison de la richesse de leur contenu, les prises de vue sont une excellente source d'information et de référence. À cet égard, elles sont fréquemment utilisées pour corroborer des enquêtes de terrain à grande échelle. Le fait que les agences de cartographie proposent des orthophotos comme produits standard confère une valeur supplémentaire à cette source d'information majeure.

Dans la mesure où les photographies aériennes à grande échelle ne reflètent que de petites portions de la surface terrestre, de nombreux clichés sont nécessaires pour couvrir de vastes superficies. Ce procédé s'avère coûteux.

L'arrivée de la nouvelle génération des capteurs haute résolution (comme IKONOS) gomment progressivement les différences, en termes de résolutions spatiale et spectrale, entre les photographies aériennes et les images satellitaires. Pour l'utilisateur, c'est l'assurance de disposer de données de haute qualité.

4.3. Sondages – Méthodes de sondage aréolaire

Les sondages à base aréolaire constituent une autre méthode d'acquisition de données sur l'occupation ou l'utilisation des sols dans le secteur agricole (rendements, etc.). Contrairement aux études par télédétection, qui couvrent la totalité de la zone concernée, le sondage à base aréolaire repose sur la sélection et l'observation d'échantillons représentatifs. L'objectif de l'échantillonnage est d'aboutir à une généralisation pertinente sans étudier l'intégralité de la zone en question.

La méthode de sondage aréolaire remplit un objectif spécifique (déterminer un ensemble de variables à étudier) et elle est couramment employée pour les statistiques agricoles (exemple : FAO 1996 ; FAO 1998 ; COTTER & NEALON 1987).

4.3.1. Principes

Le sondage aréolaire peut être considéré comme une approche statistique destinée à collecter des données pour une référence ou une unité de restitution prédéfinie. À partir de paramètres d'échantillonnage, elle se propose d'induire des valeurs représentatives de la totalité de la zone considérée. De nombreuses règles statistiques ont été élaborées pour appliquer les résultats de l'échantillonnage à l'ensemble de la population.

Le sondage à base aréolaire est défini comme un sondage dans lequel les terres sont, au moins à l'étape finale, des unités d'échantillonnage (FAO 1996). Les unités d'un sondage aréolaire sont directement liées à la zone géographique (GALLEGO 1995).

À l'intention des néophytes, voici un aperçu des principales caractéristiques du sondage à base aréolaire :

Unités d'échantillonnage

La base aréolaire de sondage consiste au départ à diviser la terre en unités (unités primaires d'échantillonnage), à partir desquelles seront extraits des échantillons représentatifs. Il existe trois principaux types d'unités :

- Les points

  Le point est l'unité d'échantillonnage la plus simple (point sur une carte ou une photographie aérienne). Il peut être localisé précisément à l'aide de ses coordonnées géographiques.

- Les lignes

  Les lignes peuvent être utilisées comme unités d'échantillonnage pour sélectionner des éléments le long d'un transect, représenté sous forme de ligne sur une carte ou une photographie aérienne.

- Les carrés ou polygones

  Couramment utilisés par les écologistes sur le terrain, les carrés ou les polygones conviennent surtout à l'échantillonnage géographique de la végétation. Contrairement aux points et aux lignes, ils sont difficiles à manipuler car leurs limites exactes ne sont pas toujours aisément identifiables. Les photographies aériennes aident à surmonter cette difficulté. Plusieurs formes, qui ont chacune leurs avantages et leurs inconvénients, sont envisageables : carré régulier ou irrégulier, avec ou sans délimitations physiques reconnaissables, etc.

Les échantillons sélectionnés sont étudiés, les variables correspondant à des informations intéressantes (hectares de cultures ou occupation/utilisation des sols) sont collectées, puis assignées à l'ensemble de la superficie ou de l'unité de référence.

Plan de sondage ou d'échantillonnage

Il existe plusieurs procédures pour sélectionner les unités d'échantillonnage :

- Échantillonnage aléatoire simple
- Échantillonnage systématique
- Échantillonnage stratifié

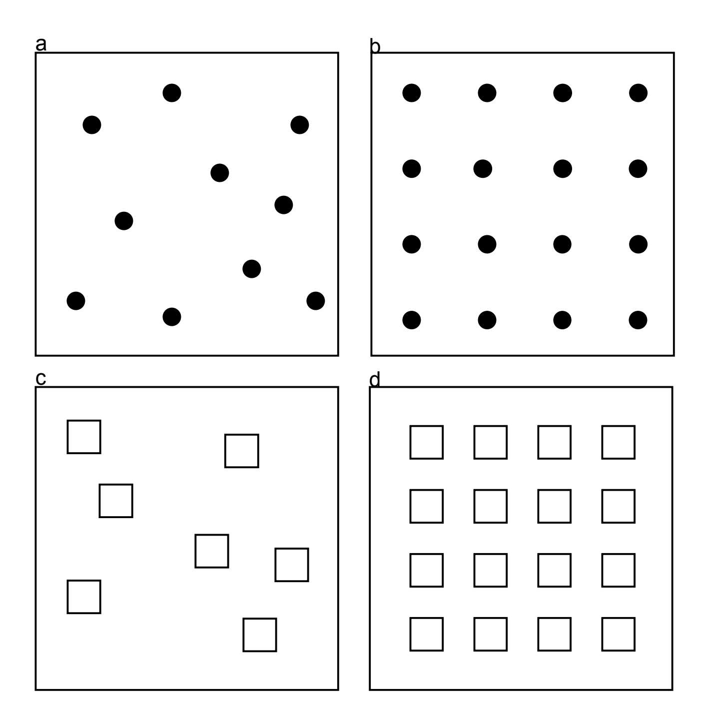


Figure 4.11 : Exemples d'unités d'échantillonnage (a,b : points ; c,d : carrés), échantillonnage aléatoire (a,c) et systématique (b,d)

Échantillonnage aléatoire simple

L'échantillonnage aléatoire est la méthode la plus simple de sélection des unités à observer. Un échantillon aléatoire simple est sélectionné par un processus qui donne à tout échantillon la même chance de sélection (sondage probabiliste). Les figures ci-dessus illustrent la subdivision d'une zone en unités d'échantillonnage (mailles) à partir desquelles un échantillon est extrait et étudié.

Échantillonnage systématique

L'échantillonnage systématique est une méthode dans laquelle seule la première unité fait l'objet d'une sélection aléatoire ou déterminée par des conditions spécifiques. Les suivantes sont extraites à intervalles réguliers.

Échantillonnage stratifié

L'échantillonnage stratifié est une méthode plus avancée de sélection des échantillons. Elle est particulièrement utile si les phénomènes étudiés sont répartis d'une manière irrégulière ou si la zone concernée est très hétérogène. La zone d'étude est divisée en plusieurs sous-zones prédéfinies (strates) qui présentent une homogénéité au regard de la distribution spatiale des caractéristiques et des attributs pertinents. Des échantillons indépendants sont sélectionnés dans chaque strate, soit au hasard soit sur un mode systématique.

La plupart des sondages à base aréolaire opérationnels font appel à la stratification afin d'augmenter la précision statistique et d'adapter le plan d'échantillonnage à la situation locale. L'homogénéisation est réalisée en fonction des connaissances *a priori* des attributs spécifiques, et en tenant compte de la distribution géographique. Dans les enquêtes agricoles, ces **strates** peuvent correspondre à certaines caractéristiques d'occupation/utilisation des sols. En principe, les images satellitaires ou les photographies aériennes facilitent la délimitation des strates et des segments.

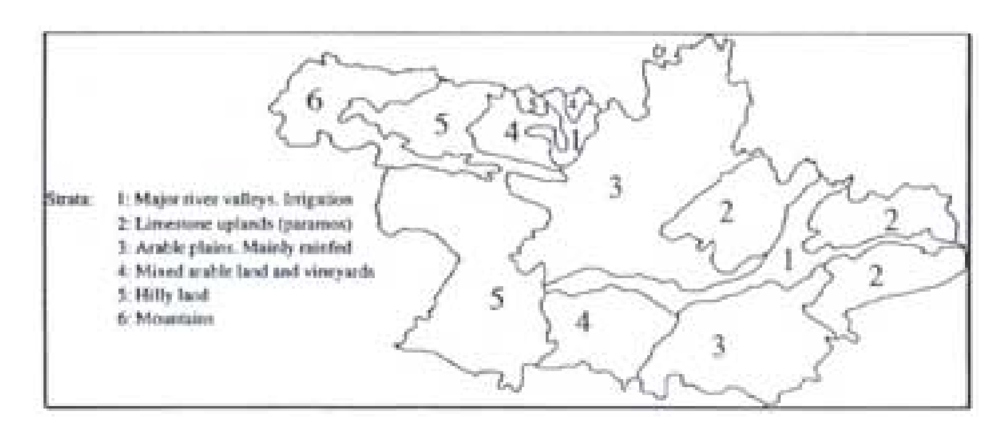


Figure 4.12 : Exemple de stratification territoriale basée sur l'utilisation des sols et le relief du terrain (TAYLOR et. al 1997)

Les unités d'échantillonnage (points, lignes, polygones) et la procédure de sélection (aléatoire, systématique, stratifiée) constituent les approches de base. Il n'existe aucune règle stricte ou standardisée pour définir et sélectionner les échantillons. Dans la pratique, le plan d'échantillonnage est déterminé par les variables à observer, la précision statistique requise, les ressources humaines et financières disponibles, et les délais d'obtention des résultats. Les approches adoptées représentent un compromis entre ces différents aspects.

Taille des échantillons

Le nombre des unités d'échantillonnage à observer est crucial dans le plan de sondage. La taille de l'échantillon détermine la précision requise pour la caractéristique de la population estimée. Son choix dépend aussi des spécificités de la superficie en question, ainsi que des ressources humaines et financières disponibles pour réaliser l'enquête.
On peut considérer que la précision de l'estimation est proportionnelle à la taille de l'échantillon.

Résultats

Contrairement à l'approche cartographique, un plan d'échantillonnage fournit uniquement des statistiques sur une région spécifique. Pour chaque plan, des règles statistiques spéciales d'extrapolation permettent d'appliquer les résultats de l'échantillon à la totalité de la zone d'étude (facteur d'expansion). Tout l'art de l'échantillonnage réside dans la capacité à sélectionner la taille d'échantillon minimale qui apportera une réponse fiable et crédible. Le « meilleur » estimateur est un estimateur sans biais dont la variance est la plus fiable.

La qualité ou la précision des estimations peut être évaluée par différentes mesures statistiques, telle que l'erreur type de la moyenne. Un large éventail de mesures statistiques a été développé pour les différentes méthodes d'échantillonnage, afin d'évaluer les résultats et d'améliorer progressivement le plan de sondage.

4.3.2. Avantages et inconvénients

Les enquêtes par sondage aréolaire présentent un avantage dans le sens elles limitent l'observation à certaines portions d'un territoire, permettant ainsi de collecter des informations très détaillées et spécifiques, qui ne peuvent être obtenues par télédétection. Cette méthode est surtout employée pour estimer des rendements agricoles mais aussi à des fins écologiques[1].

L'avantage d'un sondage à base aréolaire est la relative simplicité de sa réalisation. Il permet des retours périodiques (annuels ou saisonniers) et fiables de l'information.

Par ailleurs, des estimations de précision peuvent juger de la fiabilité et de l'efficacité.

Du point de vue conceptuel, la base de sondage aréolaire génère des statistiques valables pour les prises de décision politique d'ordre général. En revanche, elle présente relativement peu d'intérêt pour les programmes concrets (aménagement du territoire, par exemple) au niveau local, qui requièrent des données cartographiques exhaustives.

Le retour d'information des sondages aréolaires renvoie uniquement à des unités de restitution prédéfinies, généralement des divisions administratives. Si un regroupement spatial au niveau supérieur est envisageable, la désagrégation en unités plus petites est impossible. Cette contrainte n'existe pas avec les cartes exhaustives, où les données peuvent être regroupées selon diverses délimitations (découpages administratifs, unités de paysage, lignes de partage des eaux, etc.).

[1] Sondage aréolaire sur l'environnement en Allemagne, étude de l'espace rural au Royaume-Uni.

4.4. Données administratives

4.4.1. Intégration de données à partir de registres administratifs et statistiques

Au regard de la définition de « l'utilisation des sols » donnée dans l'introduction de ce manuel, il apparaît clairement que l'imagerie satellitaire ou la photographie aérienne ne peut observer qu'un nombre limité de paramètres.

Les statistiques devront prendre en compte les aspects fonctionnels et décrire la dimension socio-économique des zones concernées. Ces informations peuvent être compilées au moyen d'enquêtes ou de recensements et/ou grâce aux données déjà réunies dans les registres administratifs et statistiques, notamment sur les zones urbanisées.

La possibilité d'exploiter des données administratives en statistique dépend surtout de la finalité première des registres. Leur rôle doit être suffisamment important pour garantir une haute qualité en termes de précision et de réalité. Par ailleurs, les informations sur l'utilisation des sols contenues dans les registres doivent être pertinentes dans ce domaine spécifique. À titre d'exemple, les registres du cadastre recueillent des informations sur l'utilisation des sols, mais là n'est pas leur objet principal. Les aspects relatifs à la propriété seront toujours actualisés, mais pas ceux concernant l'utilisation des sols. Autre exemple : le SGIC (Système Intégré de Gestion et de Contrôle). Les informations sur l'utilisation des sols y sont tout à fait fiables, mais elles ne portent pas sur la population statistique des agriculteurs dans son ensemble, se limitant uniquement à ceux qui ont demandé des aides.

L'arrivée des systèmes d'information géographique (SIG) et des données géoréférencées contenues dans les registres a constitué de nouvelles et importantes sources d'information, exploitables dans les statistiques sur l'utilisation des sols. Les registres concernant l'activité économique et les informations techniques sur les propriétés immobilières et foncières peuvent être utilisés seuls ou combinés aux observations tirées des images satellitaires/orthophotos, afin de former une source solide et exhaustive de statistiques relatives à l'utilisation des sols.

4.4.2. Principes

Le premier principe consiste à établir un lien entre les entités physiques du sol et l'activité socio-économique dominante qui lui a conféré ses caractéristiques. Pour ce faire, il suffit de relier les données géocodées et géoréférencées contenues dans les registres des exploitations agricoles, les registres fonciers ou les répertoires d'entreprises (type d'activité, nombre de salariés, chiffre d'affaires, etc.) aux informations géoréférencées fournies par le cadastre (type et taille des bâtiments et des propriétés foncières, entre autres).

L'adresse ou le code cadastral de la propriété foncière sera en général la principale passerelle entre les différents registres, de sorte que les informations sur l'activité économique, par exemple, puissent être associées aux propriétés foncières ou immobilières avec des coordonnées. En l'absence d'utilisation harmonisée ou de qualité des adresses, il est possible de recourir à des approximations pour le géocodage, tels que des quadrillages fins, des secteurs de recensement/unités statistiques de base ou des portions de routes.

Il est indispensable de déterminer clairement la finalité des statistiques dérivées. Des données géoréférencées moins précises sont parfois suffisantes et offrent une solution rentable.

Des superficies plus vastes, où prédomine un type d'utilisation des sols spécifique, peuvent être regroupées dans des catégories (zones industrielles, résidentielles, etc.) au sein d'un SIG.

Les surfaces classifiées par type d'utilisation des sols peuvent être combinées et associées dans des modèles, par l'intermédiaire de bases de données numériques relatives aux voies ferrées et/ou routières, afin que les analyses du réseau permettent de suivre les interactions entre les entités.

4.4.3. Avantages et inconvénients

Cette méthode de collecte des données sur l’utilisation des sols présente des avantages en raison du lien direct entre les statistiques relatives à l’activité humaine et le sol sur lequel ces activités sont exercées. Une relation peut être établie entre la classification de l’utilisation des sols et la classification des activités CTCI (Classification Type pour le Commerce International) ou NACE (Nomenclature des Activités dans la Communauté Européenne).

Dans la mesure où le volume d’informations géoréférencées augmente, il est possible de relier les paramètres démographiques et microéconomiques directement à l’utilisation des sols. Cette approche permet de créer des bases de données globales pour la surveillance et l’analyse du statut et de l’évolution de l’utilisation des sols, et de fournir des informations contextuelles importantes sur les potentiels existants.

Cette méthode d’agrégation des statistiques relatives à l’utilisation des sols est rentable et la mise en œuvre du SIG confère une nouvelle dimension à l’exploitation des registres administratifs et statistiques. Constamment mis à jour, ces registres sont des sources fiables qui reflètent le dynamisme des zones urbanisées.

L’inconvénient tient au fait que les registres n’offrent pas toujours une solution optimale en termes d’exhaustivité, de qualité et d’actualité des données. Par ailleurs, ils ne couvrent pas toutes les zones qui devraient être répertoriées comme urbanisées. Il est généralement difficile d’y trouver des informations géoréférencées sur les espaces verts en zone urbaine, ainsi que sur les vastes superficies macadamisées mais non construites telles que certaines parties de zones industrielles, les parkings et les grands ports. Il est possible de pallier en partie cette insuffisance en utilisant des cartes topographiques à grande échelle et des relevés de cadastre.

Certains registres ne sont pas toujours exploitables car le processus de géoréférencement peut échouer. L’alignement des registres renvoie des informations non concordantes qui peuvent s’avérer extrêmement onéreuses lorsqu’elles sont traitées manuellement.

La généralisation, c’est-à-dire l’extrapolation d’informations à partir d’un site comportant un seul bâtiment et leur application à des superficies plus vastes dont l’utilisation est quasi-homogène, repose sur un ensemble de critères et de conditions préalables qui sont toujours discutables. En outre, ce processus s’accompagne de la perte de certaines informations.

5. BESOINS D’INFORMATIONS – SPECIFICATIONS DE DONNEES

5. BESOINS D'INFORMATIONS – SPECIFICATIONS DE DONNEES

L'élaboration de politiques et le processus décisionnel reposent sur la détention d'informations. Les décisions sont d'autant plus réalistes et utiles que la qualité du contenu informatif est adaptée et suffisamment complète au regard des détails thématiques et de la couverture géographique. En règle générale, politiques et décideurs n'expriment pas leurs besoins d'information en termes de spécifications de données. Dans le meilleur des cas, ils décrivent le problème et les informations nécessaires pour formuler une politique et justifier certaines décisions. Il incombe au statisticien de **traduire**/d'interpréter ce « besoin d'information » en spécifications concrètes, puis d'utiliser le système d'information statistique pour fournir les chiffres appropriés. En revanche, il lui est extrêmement difficile d'évaluer à l'avance la demande potentielle d'information, son importance et les priorités des utilisateurs.

Afin de faciliter ce processus crucial, il est absolument indispensable d'instaurer **un dialogue** entre les politiques/décideurs et les statisticiens. Ces derniers doivent posséder une **compétence multidisciplinaire** pour comprendre les problèmes dans divers domaines, identifier les indicateurs appropriés, et **définir** les **spécifications** des **données** nécessaires pour informer ces indicateurs.

À titre d'exemple, les décideurs européens doivent surveiller l'impact de la Politique agricole commune (PAC) afin de décider de sa poursuite ou de sa modification. Au cours des discussions, le problème sera spécifié plus en détail. Si une analyse de la situation économique des agriculteurs est nécessaire pour justifier le versement d'aides directes au revenu, des statistiques sur le revenu net des agriculteurs, l'emploi des membres de la famille, etc., peuvent présenter un intérêt. En cas de problèmes écologiques (pollution des eaux souterraines, par exemple), des statistiques sur les engrais utilisés, les types de terrains et les catégories d'utilisation et d'occupation des sols seront utiles. Le statisticien doit pouvoir trouver les données thématiques pertinentes, au niveau géographique approprié, pour informer ces indicateurs dans le système d'information statistique.

5.1. Quel est le « niveau » d'information requis ?

Il convient d'opérer une distinction entre les différents niveaux de politique (régionale, nationale et internationale) et le type des décisions à prendre (actions concrètes ou élaboration de cadres et programmes spécifiques). Dans les deux cas, les exigences sont très précises concernant deux aspects fondamentaux : l'aspect **thématique** et l'aspect **géographique**[2].

Du **point de vue thématique**, le niveau d'information se réfère en principe au **degré de détail**.

Par exemple, des informations de « bas » niveau sur les catégories d'occupation des sols établissent simplement une distinction entre « zone urbanisée » et « zone non urbanisée ». Côté utilisation des sols, elles se contenteraient de faire référence à une « zone artificielle », une « zone agricole utilisée » ou une « zone forestière ». À ce faible niveau détail, les spécifications des données sont beaucoup plus simples à décrire dans le cas de *l'occupation* que dans celui de *l'utilisation des sols*.

Les informations de haut niveau correspondraient, par exemple, à une classification très détaillée et désagrégée de l'occupation des sols en agriculture.

D'un point de vue thématique, des données très désagrégées sont la base d'un processus décisionnel raisonnable, tant en politique internationale que locale. La notion de « niveau européen » ne signifie pas systématiquement que le niveau de détail requis est bas. Au contraire, le niveau de détail thématique est assez élevé ; par exemple, terres en jachère (terres non cultivées donnant droit à des aides financières et éventuellement utilisées pour des installations d'énergie renouvelable, par exemple – haut niveau de détail pour la catégorie et l'espace – mise en jachère pendant une certaine période – haut niveau de détail temporel).

---

[2] La notion de « niveau » en statistique ne doit pas être confondue avec celle d'« échelle » en cartographie.

Le tableau suivant présente la nomenclature de CLUSTERS (Classification of Land Use Statistics - Eurostat Télédétection), une nomenclature pilote élaborée par Eurostat dans le cadre de son programme sur la télédétection et les statistiques (Eurostat 1996).

Tableau 5.1 : Exemple de niveau de détail thématique : CLUSTERS – nomenclature pilote pour les statistiques de l'utilisation des sols (Eurostat 1996)

	Niveau I		Niveau II		Niveau III		Niveau IV
A	Zones-artificielles	A1	Zones résidentielles Services publics	A11	Zones résidentielles	A111	Zones résidentielles continues et denses
						A112	Zones résidentielles continues moyennement denses
						A113	Zones résidentielles discontinues moyennement denses
						A114	Zones résidentielles isolées
						A115	Zones résidentielles collectives
				A12	Services publics, collectivités	A120	Services publics et collectivités
		A2	Activités industrielles et commerciales	A20	Activités industrielles et commerciales	A201	Industrie lourde
						A202	Industrie de transformation
						A203	Activités et services commerciaux et financiers
						A204	Exploitations agricoles
		A3	Infrastructures techniques et de transport	A31	Infrastructures techniques	A311	Réseaux techniques, structures de production
						A312	Installations de traitement des eaux et des déchets
				A32	Infrastructures de transport	A321	Infrastructures de transport routier
						A322	Réseaux ferrés
						A323	Aéroports et aérodromes
						A324	Installations de transport fluvial et maritime
		A4	Industries extractives, chantiers, décharges et friches industrielles	A41	Industries extractives	A410	Industries extractives
				A42	Chantiers, décharges et friches industrielles	A421	Chantiers
						A422	Décharges
						A423	Friches industrielles
		A5	Terrains aménagés à des fins de loisirs	A50	Terrains aménagés à des fins de loisirs	A501	Sites culturels
						A502	Équipements sportifs
						A503	Espaces verts ou de loisirs
B	Zones agricoles utilisées	B1	Terres arables et jachères	B11	Céréales	B110	Céréales
				B12	Plantes sarclées et plantes industrielles	B121	Plantes sarclées
						B122	Plantes industrielles non permanentes
				B13	Légumes, fleurs	B131	Légumes secs
						B132	Légumes frais
						B133	Floriculture
				B14	Jachères y compris engrais verts	B140	Jachères y compris engrais verts
		B2	Zones herbeuses utilisées à des fins agricoles	B21	Prairies temporaires et artificielles	B210	Prairies temporaires et artificielles
				B22	Prairies, pâturages permanents	B220	Prairies et pâturages permanents
				B23	Pâturages pauvres	B230	Pâturages pauvres
		B3	Cultures permanentes	B31	Arbres fruitiers et baies	B310	Arbres fruitiers et baies
				B32	Agrumeraies	B320	Agrumeraies
				B33	Oliveraies	B330	Oliveraies
				B34	Vignes	B340	Vignes
				B35	Pépinières	B350	Pépinières
				B36	Cultures industrielles permanentes	B360	Cultures industrielles permanentes
C	Zones forestières	C1	Zones forestières boisées	C11	Feuillus caducifoliés	C110	Feuillus caducifoliés
				C12	Feuillus sclérophylles	C120	Feuillus sclérophylles
				C13	Résineux	C130	Résineux
				C14	Lignicultures	C140	Lignicultures
		C2	Zones forestières non boisées	C21	Coupes rases	C210	Coupes rases
				C22	Autres zones forestières improductives	C220	Autres zones forestières improductives
D	Zones arbustives ou herbeuses	D1	Végétation arbustive	D10	Végétation arbustive	D101	Formations arbustives des régions tempérées, montagnardes ou arctiques
						D102	Formations arbustives xérophytiques
		D2	Végétation herbacée	D20	Végétation herbacée	D201	Prairies et pelouses des régions tempérées, montagnardes ou arctiques
						D202	Steppes et prairies arides
E	Zones avec peu ou pas de végétation	E0	Zones avec peu ou pas de végétation	E01	Sol nu	E011	Rochers et éboulis
						E012	Dunes et plages
				E02	Glaciers et, neiges éternelles	E020	Glaciers et neiges éternelles
				E03	Zones incendiées	E030	Zones incendiées
F	Zones humides	F1	Zones humides	F10	Zones humides	F101	Marais et marécages
						F102	Tourbières

	ou en eau					F103	Autres zones humides
		F2	Eaux intérieures	F20	Eaux intérieures	F201	Cours d'eau et plans d'eau intérieurs
						F202	Plans d'eau intérieurs utilisés pour la pisciculture
		F3	Eaux côtières	F30	Eaux côtières	F301	Estuaires, lagunes
						F302	Plans d'eau côtiers utilisés pour les cultures marines

Du **point de vue géographique**, le niveau des informations statistiques fait référence au découpage géographique (unité de restitution dans un système de référence statistique régionale, par exemple). L'unité de restitution est celle dans laquelle les données statistiques sont représentées. Elles peuvent avoir plusieurs dimensions temporelles (intervalles) et spatiales (unités territoriales). Dans ce dernier cas, les données sont géocodées, c'est-à-dire qu'elles ont une référence géographique (une ville ou une localité spécifique, une région ou un pays, ou encore un groupe de pays comme l'Union européenne). Il est prévu que les informations statistiques soient **harmonisées et comparables** sur l'ensemble de la couverture géographique du système de référence régionale. Au niveau européen, le système de référence utilisé est la nomenclature des unités territoriales statistiques (NUTS[3]).

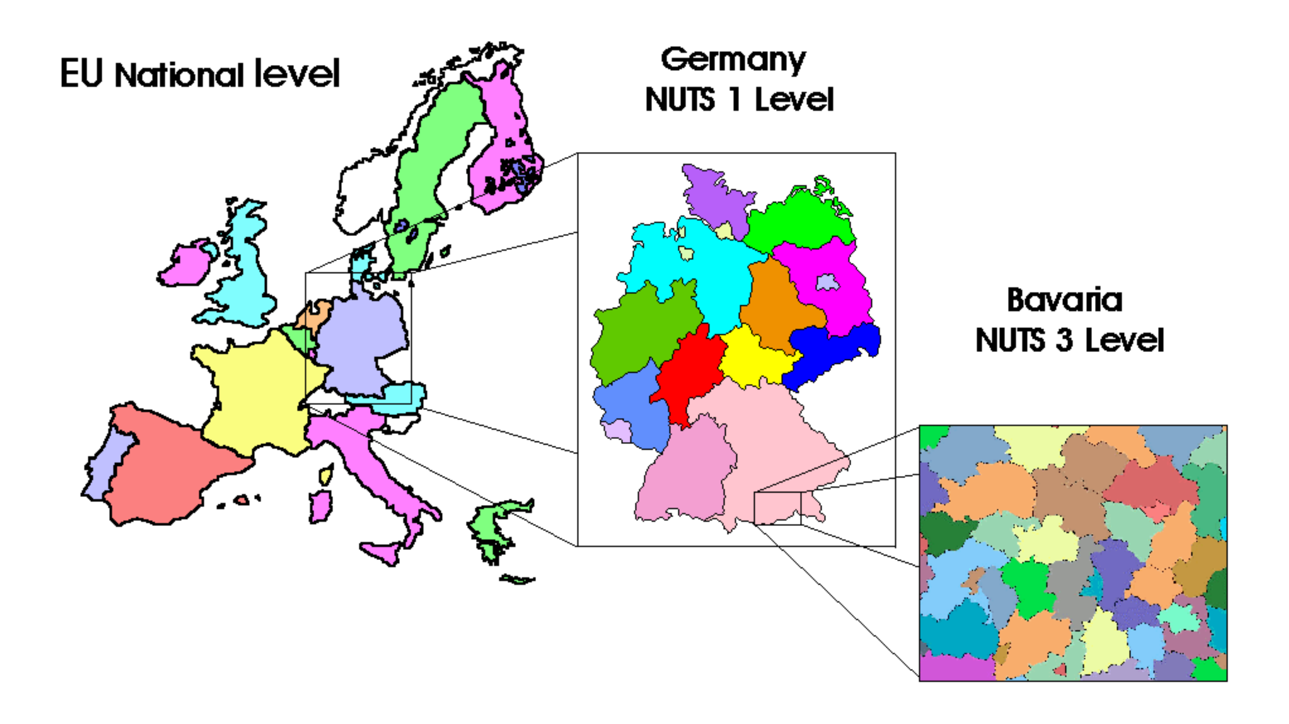


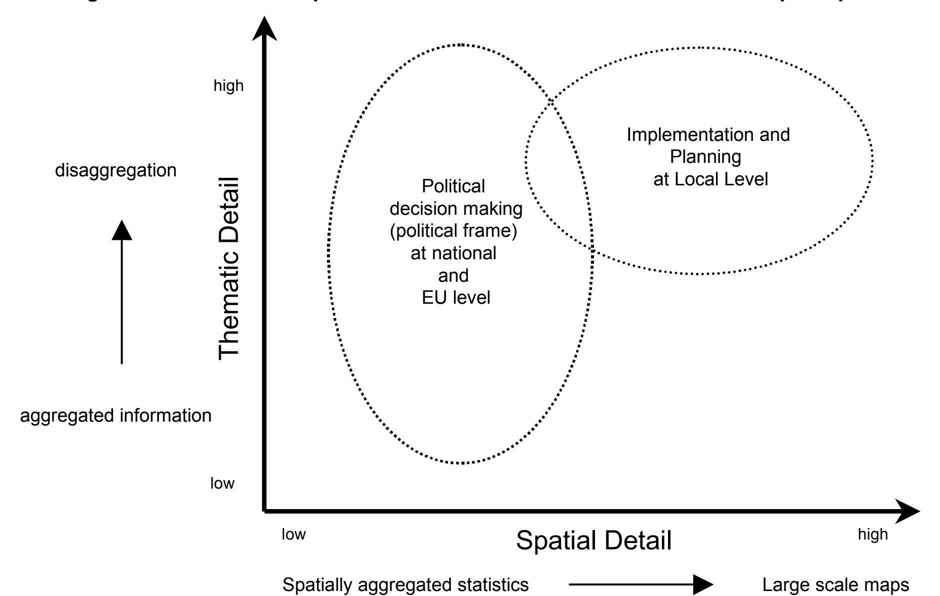
Figure 5.1 : Exemple de niveau de détail spatial : niveau NUTS 0 = États membres de l'UE ; NUTS 1 (Allemagne) = Länder ; NUTS 3 (Bavière) = Kreise

En ce qui concerne la dimension spatiale, il convient de faire la distinction entre politique générale et programme d'actions concrètes. Aux niveaux international et national, l'agrégation des statistiques sur le plan spatial ne suffit pas. Au niveau local, en particulier dans le cadre d'une programme d'action, une désagrégation spatiale poussée est requise ; en effet, plus la politique touche aux préoccupations locales, plus l'approche cartographique devient importante.

[3] Dans le cadre des statistiques au "niveau européen", il est nécessaire que ces chiffres soient harmonisés (comparables) entre les différents États membres, à une certain niveau de découpage territorial (niveaux NUTS : niveau « européen » = données nationales = NUTS 0 = 15 États membres ; NUTS 1 = 77 régions ; NUTS 2 = 206 provinces ; NUTS 3 = 1031 comtés ; NUTS 4 = 1074 districts ; NUTS 5 = 98 433 communes).

Le graphique ci-dessous essaie d'illustrer et de résumer la tendance générale : l'agrégation des statistiques au niveau spatial est d'autant plus importante que le niveau politique est élevé.

Figure 5.2 : Besoins et spécifications de données à différents niveaux politiques



5.2. Systèmes d'information polyvalents

Les systèmes d'information statistique doivent fournir les **données source (de base)** répondant aux besoins d'information des utilisateurs dans différents domaines thématiques. Un nombre limité de données de base, dont les spécifications temporelles et spatiales sont cohérentes, peut être considéré comme une source flexible, exploitable dans divers domaines.

À titre d'exemple, de nombreux indicateurs statistiques économiques et démographiques utilisés par les décideurs reposent sur des données de base (population, âge, niveau d'étude, revenus, etc.). Cet usage polyvalent améliore la rentabilité du système statistique, puisque de nombreux utilisateurs contribuent à la collecte, au traitement, à la diffusion des données, etc. et, par conséquent, au coût général du système.

Malheureusement, la situation est quelque peu différente en ce qui concerne les statistiques sur l'occupation et l'utilisation des sols. Il n'existe aucune base de données polyvalente au niveau européen ; plusieurs systèmes d'information ont été mis en place pour répondre à des demandes précises dans des domaines spécifiques (agriculture, environnement, etc.). En Europe, plus de 70 systèmes ont trait à l'utilisation et à l'occupation des sols (CROI 1999).

S'ils poursuivent tous le même objectif – fournir des informations sur l'utilisation et l'occupation des sols – leurs contenus sont si différents qu'ils sont rarement comparables. La situation empire quand ces informations sont exploitées dans d'autres domaines thématiques comme l'environnement.

Voici un bref récapitulatif (tiré de JANSEN, L.J.M. ; A. DI GREGORIO (1999)) des causes et des conséquences de cette situation peu satisfaisante :

- Approche sectorielle correspondant aux conditions territoriales spécifiques

  La plupart des systèmes d'information suivent une approche sectorielle qui répond aux besoins spécifiques des utilisateurs et aux conditions locales.

  Par conséquent, ils ne permettent pas une définition exhaustive des classes d'occupation des sols.

L'échelle et la méthode de collecte des données (la télédétection, par exemple) sont souvent choisies par rapport à un objectif précis. Dans certains cas (les classes d'occupation des sols basées sur les données NOAA AVHRR), les classes constituées peuvent être étroitement tributaires des modes d'acquisition des données.

- Cohérence

  La plupart des systèmes d'information sur l'occupation/utilisation des sols pâtissent de leurs incohérences internes.

  Les critères utilisés pour constituer des classes sont rarement appliqués systématiquement, ou bien les définitions de classes sont imprécises, ambiguës ou inexistantes. Les critères de diagnostic retenus, ainsi que leur combinaison, vont souvent à l'encontre d'une délimitation précise des classes. Pourtant, il s'agit là d'une exigence de base de n'importe quel système.

  Il est rare que les combinaisons des différents éléments de diagnostic décrivant une classe soient examinées dans leur ensemble ; la mise en œuvre de toutes les combinaisons de classificateurs possibles génèrerait en effet une multitude de classes, que les méthodes actuelles de description sont incapables de gérer. Par conséquent, les systèmes existants présentent souvent des lacunes dans l'application systématique des critères de diagnostic.

  Ces incohérences se soldent par une imprécision dans la définition des classes entre plusieurs classifications, par des recoupements ou des lacunes.

  Les systèmes d'information sont parfois de simples légendes plutôt qu'une classification systématique : la description porte uniquement sur une portion de l'éventail des classes possibles. L'utilisateur ne peut se référer à un système de classification qui ne permet pas d'établir des comparaisons avec d'autres systèmes.

Ces inconvénients traduisent le manque de compatibilité entre les systèmes de classification, ou entre les classifications et les légendes. Quelle que soit l'adéquation des systèmes existants à leurs tâches spécifiques, ils empêchent un large public d'exploiter le résultat de ces classifications dans de nombreuses applications.

Le développement de systèmes d'information polyvalents devrait être encouragé afin de mettre un terme à cette situation peu satisfaisante.

Un système d'information polyvalent sur l'utilisation et l'occupation des sols poursuit plusieurs objectifs. Les données extraites doivent fournir :

- des informations comparables extraites de données harmonisées ;
- des informations génériques exploitables dans plusieurs domaines thématiques à des fins spécifiques, afin que différents utilisateurs puissent se référer à la même source d'informations.

Les avantages d'un tel système résident dans ses capacités de synergie. Quand les ressources sont limitées, les efforts déployés pour construire un système d'information unique et harmonisé sur l'occupation et l'utilisation des sols permettent d'économiser de l'argent, du temps et des ressources humaines.

6. GESTION ET MANIPULATION DES DONNEES

6. GESTION ET MANIPULATION DES DONNEES

6.1. Généralisation et agrégation

6.1.1 Généralisation

La généralisation est la « réduction des détails ou la simplification de la réalité ».

En *CARTOGRAPHIE*, la généralisation porte sur les caractéristiques géométriques et thématiques du contenu cartographique. Les cartes sont des représentations de la réalité et, à ce titre, elles ne peuvent contenir tous les détails, au risque sinon de perdre en lisibilité. Le **processus de généralisation** comprend les phases de simplification, sélection et classification.

Simplification

Tous les éléments possèdent trois dimensions intrinsèques : longueur, largeur et hauteur. Les cartes sur papier sont bidimensionnelles, la hauteur (z) étant perdue. Les éléments géographiques peuvent être simplifiés comme suit :

Surface	2D	Les polygones peuvent être remplis pour les distinguer des lignes (ex : les lacs)
Ligne	1D	Les arcs peuvent être ornés de motifs pour les distinguer des délimitations (ex : les routes)
Point	0D	Symboles ponctuels (ex : les points de repère)

L'une des méthodes de simplification est la RÉDUCTION : au fur et à mesure que l'échelle est réduite (du 1/50 000e au 1/ 500 000e, par exemple), les éléments perdent de leur précision :

Surfaces -> Lignes (ex : les rivières, les routes présentées à grande échelle comme des superficies, symbolisées par deux lignes, sont réduites à une seule ligne)

Surfaces -> Points (ex : les contours des villes sont réduits à un point)

Lignes -> Points (ex : les ponts)

Le niveau de détail des lignes et des surfaces, ainsi que le nombre total de points, lignes et surfaces diminuent également.

Sélection

Un autre mode de généralisation consiste à sélectionner des éléments selon la finalité de la carte (togrographique, thématique). Le nombre des éléments conservés dépend surtout, en dehors du thème, de l'échelle : si elle est divisée par deux, la superficie totale couverte est divisée par 4, et le nombre d'éléments (points, lignes et surfaces) est réduit en conséquence.

Processus de sélection :

- Thématique : Agrégation (fusion de plusieurs éléments ; couramment utilisé pour les surfaces)

  Élimination (suppression de certains éléments)
- Géométrique : Lissage (suppression des détails relatifs aux formes et contours)

Classification

Bien que la plupart des éléments aient des caractéristiques propres (il n'existe pas deux forêts ou deux villes identiques), ils doivent être regroupés en classes pour faciliter la lecture et la représentation. On distingue trois types de classifications ou regroupements de données :

- Nominal : par type, qualitatif (essences d'arbres, oléoducs ou gazoducs, types de bâtiments, etc.)
- Ordinal : classement ordonné, impliquant la notion de taille (risque d'incendie – élevé, moyen, faible ; hiérarchie des routes – principales, secondaires, chemins)
- Intervalle : quantitatif, par taille (populations des villes, pluviométrie) – exprimé en valeur totale, en ratios, densités ou pourcentages

6.1.2. Agrégation

En *STATISTIQUE*, la généralisation ou la perte de détail est implicite dans le processus d'agrégation. Deux aspects doivent être pris en compte :

- La dimension thématique ; par exemple, combiner plusieurs variables dans un « agrégat » ou un « indicateur » (= compression de l'information) ou ajouter les chiffres sur l'occupation des sols correspondant à différents niveaux d'une classification jusqu'à un niveau supérieur (exemple, des 44 classes CORINE Land Cover du 3ème niveau jusqu'aux 15 classes du 2èmeniveau).
- La dimension spatiale ; ajouter (ou faire la moyenne) des chiffres associés à de petites unités de restitution jusqu'à atteindre des superficies plus importantes (niveaux supérieurs du système de référence régionale), sans en modifier la signification.

Les données de recensement font en principe l'objet d'une agrégation spatiale pour garantir la confidentialité des statistiques.

Il convient d'opérer une distinction majeure dans l'utilisation des cartes en statistique :

- Cartographie thématique ou statistique

  Les statistiques sont présentées sur des cartes généralisées (format numérique ou analogique). Les éléments sélectionnés, essentiellement les délimitations des régions de référence, et la toponymie sont suffisants. Cela ne signifie pas que le contenu statistique des cartes est également généralisé, c'est-à-dire agrégé. Des statistiques extrêmement précises peuvent figurer sur des cartes à petite échelle très généralisées.

- Utilisation de cartes ou de données géographiques dans les statistiques

  Quand des données géographiques sont utilisées pour préparer des enquêtes statistiques, il est essentiel de recueillir des données géométriques précises et fiables, surtout sur une grande échelle, afin d'identifier des bâtiments ou des parcelles par exemple. Le géomètre a besoin de nombreux autres éléments pour déterminer l'unité d'observation (tracé de bâtiments, limites des parcelles fixées par le cadastre, réseau de transport, points de repère, voies d'eau, etc.). Leur prise en compte affecte la qualité des statistiques obtenues.

6.2. Intégration des données

Le processus d'unification des sources de données existantes dans un cadre unique est appelé intégration de base de données. (DEVOGELE 1998)

Dans le domaine de l'occupation/utilisation des sols, les données possèdent une dimension spatiale. Des outils logiciels spécifiques associés à des systèmes de gestion de base de données (SGBD) ont

été développés pour gérer les caractéristiques géographiques : les systèmes d'information géographique (SIG). Selon *(GOODCHILD 1997)*, un SIG est :

- un système permettant l'acquisition, le stockage, la manipulation et la diffusion d'informations géographiques ;
- une classe de logiciel.

Un exemple pratique de SIG combine logiciel, matériel, données, utilisateur, etc., pour résoudre un problème, étayer une décision, contribuer à l'élaboration d'un plan.

Le SIG intègre les données sur la base de leurs références géographiques. Les informations provenant de différentes sources peuvent décrire la même zone et leur intégration prend alors en compte la dimension géographique. Principaux problèmes :

- comprendre la signification des données (sémantique) ;
- établir une correspondance entre des contenus similaires (système de classification) ;
- choisir des techniques de traitement adéquates.

L'intégration des données doit être envisagée dans le cadre du **système d'information**, en l'occurrence un système sur l'occupation et l'utilisation des sols. Celui-ci repose en principe sur les éléments suivants :

- **La classification :** organiser les groupes d'objets (classes) dans une structure (essentiellement hiérarchique) qui définit leurs relations les uns par rapport aux autres.
- **La nomenclature :** attribuer des noms et donner des descriptions détaillées des propriétés associées aux types d'occupation et d'utilisation des sols, afin de permettre l'identification de chaque entité (parcelle, biotope, bâtiment) dans les groupes d'objets (classes d'occupation et d'utilisation des sols).
- **L'identification :** affecter chaque entité à un groupe d'objets.

La mise en œuvre d'un système d'information comprend deux phases :

- **La phase de conception et de développement** des éléments susmentionnés et des spécifications, tels que le modèle des données, la structure de la base de données, la gestion de la qualité[4], les procédures de mise à jour, l'administration et la diffusion des informations ;
- **La phase de collecte des données** : la sélection des méthodologies et des techniques appropriées, l'application des règles prédéfinies aux objets à classifier, les contrôles qualité internes (l'intégration est l'une des techniques visant à générer de nouvelles informations à partir de plusieurs sources) ;
- **Le contrôle qualité** : la mise en œuvre de procédures de contrôle de la qualité des données ;
- **La diffusion des données** : l'organisation de la diffusion des données (distribution physique, support aux utilisateurs, etc.) ;
- **La mise à jour et la maintenance** : l'adoption de procédures de mise à jour (la méthodologie peut différer de celle en vigueur lors de la première phase de collecte des données), dont la gestion de la qualité et le contrôle de la base de données.

[4] La gestion de la qualité est étroitement liée à la gestion de projet et inclut le contrôle et la révision de la mise en œuvre, la création d'une base de métadonnées, etc.

De nombreuses disciplines (agriculture, sylviculture, écologie, aménagement du territoire et urbanisme, etc.) utilisent des informations sur l'occupation ou l'utilisation des sols qui répondent à des paramètres spécifiques, selon l'objectif poursuivi. Il existe pratiquement autant de bases de données que de projets faisant appel à des informations de ce type (plus de 70 en Europe, au niveau national ou international (CROI 1999)). Ces « systèmes d'information » contiennent des données difficilement comparables et compatibles, d'où une situation de faible rentabilité.

L'idéal serait de recueillir, en partant de zéro et par le biais d'enquêtes sur le terrain, des informations de base avec un très haut degré de précision, comme il a été proposé pour le système d'information topo-cartographique officiel allemand ATKIS (*Amtliches Topographisch-Kartographisches Informationssystem*). En raison de son coût extrêmement élevé et du temps nécessaire à sa réalisation, les autorités allemandes ont décidé de numériser les cartes topographiques existantes (au 1/25 000e)[5].

L'intégration de données existantes provenant de sources multiples est une autre solution pour améliorer la rentabilité. Le choix de la mettre en œuvre pour générer une nouvelle base de données dépend fortement de la finalité et des objectifs. Il convient de les analyser clairement et d'évaluer les conséquences de l'intégration sur la qualité et la fiabilité des données obtenues.

6.2.1. Phases d'intégration des données

Le processus d'intégration des données comprend en principe les phases suivantes (DEVOGELE 1998) :

Compilation et préparation

Cette phase englobe toutes les activités nécessaires à l'intégration des diverses bases de données :

- l'identification et le regroupement des données pertinentes (y compris les métadonnées) issues de sources diverses ;
- la conversion dans un format unique (harmonisation des modèles de données hétérogènes via la création d'un modèle commun au niveau logique ; conversion des données vectorielles en données tramées, et inversement, au niveau technique) ;
- l'enrichissement des descriptions de données afin que les futurs utilisateurs puissent mesurer la qualité de la base de données intégrée : description explicite de la projection cartographique et du système de référence utilisés, les procédures de conversion appliquées, etc. ;
- la création de thésaurus au niveau global (pour toute la base de données) et local (pour chaque source).

Analyse des correspondances

L'analyse des correspondances entre catégories connexes et la détection d'éventuels conflits consistent à identifier et à détailler toutes les corrélations entre les données i) au niveau de la description (métadonnées) et ii) au niveau des objets (données). Des techniques et des outils semi-automatiques d'analyse des similitudes permettent d'établir les correspondances sur la base de la connaissance et du sens (sémantique). En outre, la collaboration avec des spécialistes des domaines source contribue à une meilleure compréhension des données et à la création de tables de conversion.

L'analyse des correspondances entre plusieurs bases de données source permet de formuler des recommandations/directives sur les possibilités d'intégration, les manipulations à effectuer à cet effet, et les contenus qui ne peuvent absolument pas être intégrés.

---

[5] La collecte de données « idéale » est réalisée parallèlement au processus régulier et continu de mise à jour des cartes de base (1/5 000e).

Harmonisation

Au niveau thématique, les catégories ou classes d'occupation et d'utilisation des sols doivent être harmonisées en fonction des résultats de l'analyse des correspondances. Cette étape du processus d'intégration est la plus délicate. Au niveau technique, les données géographiques doivent être harmonisées pour permettre une analyse synoptique (tel qu'un recouvrement) : mise en corrélation des points et des lignes (limites de polygones) provenant de différentes sources et relatifs à des objets similaires ; par exemple, le réseau hydrographique de la carte topographique est comparé à celui tiré d'une classification par image satellitaire ; les délimitations d'une terre arable extraites d'une couche de données relatives à l'utilisation des sols sont comparées à celles issues d'une couche de données relatives aux zones forestières, etc. C'est une des étapes les plus longues de la phase d'intégration.

Incorporation

L'incorporation consiste à réunir plusieurs bases de données en une seule.

À titre d'exemple, les **points** centraux de bâtiments tirés d'un registre, avec leurs attributs (= points), peuvent être intégrés aux tracés des bâtiments extraits d'une base de données à grande échelle, par exemple un plan de la ville au 1/5 000e (= lignes), pour créer une nouvelle couche de données **Surface** urbanisée (= polygones). Pour la compléter, il est possible d'incorporer d'autres données numériques sur les zones boisées, avec une échelle, une définition, un degré de précision et d'exactitude particuliers – en termes de fiabilité thématique - et recueillies à l'aide d'une méthode spécifique.

Généralisation

Une généralisation géométrique et thématique est évidemment indispensable à une intégration pertinente de données dont les spécifications et les échelles sont disparates.

Les procédures de généralisation peuvent aussi aider à préparer les données pour différentes applications. La figure 6.1 résume ce processus.

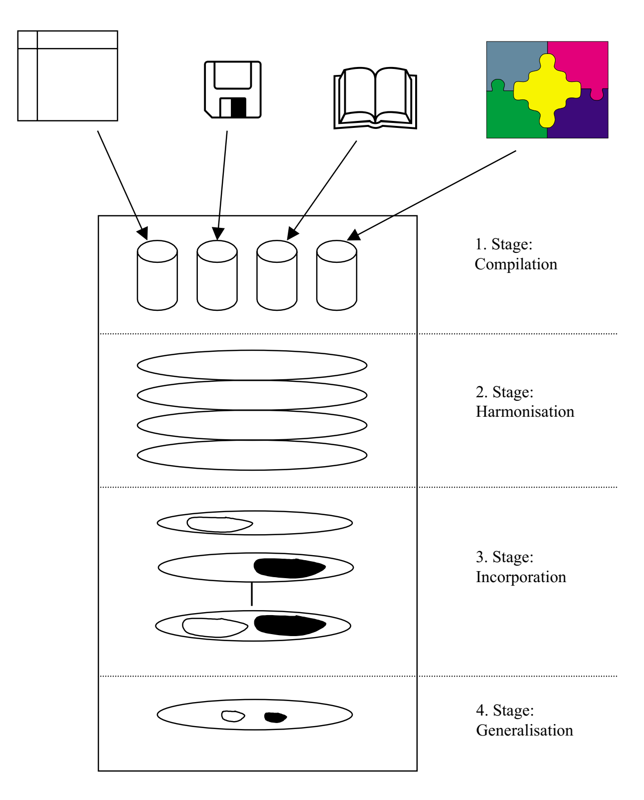


Figure 6.1 : Méthode d'intégration des données

6.2.2. Problèmes liés à l'intégration des données

Comme nous l'avons déjà indiqué, il est nécessaire de résoudre plusieurs types de problèmes relatifs aux aspects techniques du processus d'intégration, aux concepts et aux techniques d'observation mises en œuvre pour générer les bases de données source, et à la qualité du nouveau contenu.

Ces problèmes sont examinés brièvement dans les pages suivantes.

Problèmes conceptuels

Incompatibilité des paramètres système des données source

Les besoins spécifiques et l'objectif poursuivi doivent motiver la conception des composantes de base des données source, la méthode retenue pour la collecte, le choix des paramètres et des techniques appliquées.

Or, l'étude des systèmes européens a montré au contraire que nombre d'entre eux avaient été conçus en fonction de la source de données et des outils de collecte, notamment là où la télédétection est utilisée (carte d'occupation des sols de la Grande-Bretagne, CORINE Land Cover, carte forestière de l'Europe à partir de données NOAA-AVHHR, etc.).

Systèmes de classification

Les différences entre les définitions des classes des systèmes actuels n'autorisent généralement pas une comparabilité directe, ou même une compatibilité, entre les bases de données source. L'analyse des définitions détaillées des classes d'occupation et d'utilisation des sols est une étape indispensable avant toute reclassification des données. Dans le cadre de cette analyse des correspondances, il convient de prêter une attention particulière à la finalité de la nouvelle base. L'analyse devrait être confiée à une équipe multidisciplinaire compétente dans les différents domaines des données source (statisticiens agricoles, techniciens forestiers, écologistes ou urbanistes, etc.). Elle produit des règles de reclassification et des tables de conversion établissant le lien entre les nomenclatures.

Il existe plusieurs approches pour construire des passerelles ou des interfaces entre les nomenclatures : l'approche de l' ITE repose sur la création d'une nomenclature de base ; l'approche d'Eurostat se fonde sur les noyaux et les limites ; l'approche de la FAO définit des classificateurs pour l'occupation et l'utilisation des sols.

Différentes méthodes de collecte des données : quand la cartographie rencontre la statistique

L'échantillonnage statistique et la cartographie appliquent des concepts distincts pour collecter les informations sur l'occupation et l'utilisation des sols.

La cartographie, au sens propre du terme, présente l'avantage de couvrir l'intégralité du territoire considéré. L'échelle géométrique est de toute première importance pour l'application des données : une grande échelle (1/1 000e au 1/10 000e) garantit une géométrie extrêmement précise et peut aussi fournir les détails thématiques requis pour la gestion de la propriété et l'urbanisme ; des échelles plus petites (à partir du 1/50 000e) fournissent des informations généralisées en termes de précision thématique et spatiale, qui conviennent pour l'aménagement du territoire et la surveillance de l'environnement. Les inconvénients d'une approche cartographique exhaustive tiennent aux coûts élevés et aux délais nécessaires à sa mise en œuvre sur de vastes superficies.

Les méthodes d'échantillonnage sont plus économiques et, selon le mode de sondage adopté, elles fournissent rapidement des informations détaillées sur l'occupation et l'utilisation des sols. Les inconvénients sont liés à une résolution spatiale grossière en termes d'unité de restitution (la notion d'échelle est bannie dans ce contexte), selon la méthode employée et la stratification finale – c'est-à-dire que les informations reflètent une zone de référence régionale « artificielle » et non un lieu géoréférencé avec précision comme en cartographie.

Les conséquences de l'intégration de ces types de données doivent être analysées avec soin et prises en compte dans les applications. Dans certains cas, l'intégration peut être extrêmement utile (par exemple, l'inventaire forestier multi-source finlandais, TOMPPO 1992).

Problèmes liés à l'observation

Différentes techniques de capture et de traitement des données

Les données source sont généralement collectées au moyen de différentes techniques selon l'approche retenue. Dans les enquêtes statistiques, elles sont recueillies dans le cadre de sondages sur le terrain et de réponses à des questionnaires, ou par des mesures de variables. Les résultats sont étroitement liés à la stratégie d'échantillonnage et aux compétences de l'analyste. Les enquêtes sur le terrain peuvent fournir des informations très précises sur un nombre limité d'échantillons (points, segments).

La cartographie à grande échelle nécessite des enquêtes sur le terrain pour récolter des données très précises ; en revanche, pour de vastes superficies, ce sont surtout des données d'observation de la terre (photographies aériennes et imagerie satellitaire) qui sont interprétées visuellement ou numérisées pour mettre à jour ou créer des bases de données.

Les résultats de la photo-interprétation des données d'occupation et d'utilisation des sols dépendent en grande partie des compétences et de l'expérience de l'analyste, ainsi que des informations annexes dont il dispose. Il n'est pas recommandé d'exploiter ces données pour extraire des statistiques en raison du biais introduit par les erreurs systématiques non dues à l'échantillonnage (pas de compensation entre les erreurs par commission et omission) et de la taille minimale des polygones interprétés (GALLEGO & CARFAGNA 1998). Les paramètres appliqués dans les traitements automatiques ou semi-automatiques sont documentés et permettent ainsi de répéter le processus ; mais leur définition requiert la compétence d'un spécialiste des images, ce qui est, là encore, subjectif.

L'intégration des données issues des différentes méthodes de collecte a un fort impact sur la fiabilité de la base de données obtenue. Le traitement des données doit être documenté avec précision afin que l'utilisateur final puisse évaluer l'adéquation du résultat à ses besoins.

Différentes échelles et approches thématiques

Une carte est un modèle simplifié de la réalité d'une superficie donnée. Les objets sont classifiés et des symboles sont utilisés pour les représenter. Les cartes à grande échelle contiennent davantage d'objets et de symboles, avec des attributs plus détaillés, une précision et une exactitude géométrique relativement élevée. Les cartes à petite échelle présentent uniquement des informations généralisées ; le contenu et la nomenclature sont en principe moins riches.

Dans un SIG numérique, les contraintes imposées par le format du papier sont abolies. Dans l'idéal, les données enregistrées devraient être collectées dans le cadre d'enquêtes sur le terrain avec une précision et exactitude géométrique très élevées, pour ainsi dire « affranchies de toute échelle ». En raison du coût et des délais nécessaires pour couvrir de vastes superficies, les données sont actuellement numérisées à partir des cartes existantes, à des échelles variées, de sorte que la précision / l'exactitude géométrique dépend de l'échelle d'origine et de la méthode de collecte des données source.

L'intégration de données géographiques à petite échelle – dont la richesse thématique est de « faible qualité » (classes agrégées d'occupation ou d'utilisation des sols) – dans une base de données aggrave le problème des systèmes de classification (en termes de correspondance des classes) et ne contribue guère à la polyvalence des informations. Plus les données sont agrégées (= pauvres en détails), moins elles sont exploitables par d'autres traitements dans des applications polyvalentes. La plus petite échelle des bases de données source détermine la précision et l'exactitude géométrique de la base de données géographiques obtenue.

À titre d'exemple, l'imagerie satellitaire possède une résolution géométrique, correspondant à la taille des pixels, qui dépend du type de capteur, des paramètres de l'orbite de la plate-forme et du sol (altitude, énergie du relief). Les données peuvent être présentées à différentes échelles mais la résolution géométrique, de 30 m x 30 m pour les images Landsat TM, et le contenu ne sont pas modifiés. Les informations sur l'occupation des sols extraites des images satellitaires conservent cette géométrie. Lors de l'intégration de petits objets provenant d'autres sources (données géographiques ponctuelles d'un registre, par exemple), la précision du positionnement est très importante.

Le processus de généralisation peut être perçu comme une normalisation à la fois thématique et géométrique, qui permet d'intégrer des données à différentes échelles sans introduire trop de distorsion géométrique dans le résultat, tant que l'échelle des données obtenues est inférieure à la plus petite échelle des données source.

Différentes périodes de collecte des données

L'utilisation de données source recueillies à différentes dates (ou sur une courte période) conduit à des incohérences temporelles dans la base de données. Par exemple, certaines ont été collectées en 1988 pour une portion de la zone observée, et en 1994 pour le reste. Les modifications ne sont identifiables que dans la mesure où la date de chaque unité de surface est connue. Par exemple, les statistiques du territoire suisse sont tirées de la photo-interprétation des points (avec une grille d'un hectare) sur des clichés aériens couvrant toute la Suisse pour la période 1979/85 et 1992/97 (OFFICE FÉDÉRAL DE LA STATISTIQUE 1996).

Problèmes liés à la qualité et à la précision

Difficile d'évaluer la qualité de l'intégration des données. Pour permettre à l'utilisateur de mesurer la qualité d'une base de données, la principale information requise concerne les procédures de création des données, l'une des composantes des « méta-informations ». Le contrôle de la qualité doit prévoir l'analyse de la propagation des erreurs contenues dans les données source et diffusées par l'intermédiaire des processus d'intégration et de généralisation. L'évaluation de la pertinence se mesure en termes de précision géométrique et d'exactitude thématique.

Méta-informations et généalogie des données

Chaque base de données doit être accompagnée d'informations détaillées sur l'approche, la date et la méthode de collecte, les techniques adoptées pour le traitement des informations, l'évaluation du degré d'exactitude spatiale et thématique du fournisseur, etc. (HUNTING 1997). Ces méta-informations sont déterminantes dans l'intégration des données. Ce n'est que récemment que la standardisation des métadonnées a été examinée par différentes institutions, comme le Comité Européen de Normalisation (CEN) avec sa norme CEN/TC 287, l'organisation internationale de normalisation (ISO) avec ISO/TC 211 ou MEGRIN, avec son répertoire descriptif des données géographiques (*Geographical Data Description Directory*, GDDD).

Évaluation de l'exactitude thématique

L'évaluation de l'exactitude thématique est une opération complexe, mais indispensable, dans le processus de création d'une base de données. Elle est beaucoup plus délicate dans les bases intégrées, en raison des différentes sources d'erreur et du traitement des données lui-même. En ce qui concerne les statistiques sur l'occupation et l'utilisation des sols, l'exactitude peut être vérifiée à travers une double enquête menée sur un échantillon d'interviews, afin de calculer leur pertinence à l'intérieur d'un intervalle de confiance défini ou, comme c'est le cas des statistiques relatives au territoire suisse, à travers une double interprétation de chaque point par différents experts et enquêtes sur le terrain. En règle générale, l'exactitude thématique des données sur l'occupation/utilisation des sols extraites de l'imagerie satellite est évaluée sur la base de la « réalité de terrain » reflétée par les sondages aréolaires tirés d'une étude sur le terrain ou d'autres sources.

Concernant la propagation des erreurs lors de l'intégration des différents types de données, il convient de réaliser une analyse approfondie des sources d'erreur potentielles dans les bases de données spatiales : erreurs de calcul, de spécification, d'échantillonnage, de mesure ou erreurs stochastiques.

Évaluation de l'exactitude géométrique

La précision et l'exactitude géométriques dépendent de la méthode de collecte initiale et de l'échelle des données source. L'évaluation de l'exactitude géométrique doit prendre en compte le traitement des données : numérisation ou balayage, ré-échantillonnage, rectification géométrique par rapport à d'autres coordonnées géographiques et systèmes de projection, procédures de généralisation, etc.).

Une bonne méthodologie d'évaluation inclut, par exemple, des points de contrôle utilisant des mesures tirées d'études sur le terrain, de cartes topographiques ou GPS (*Global Positioning System*, système américain de navigation par satellite). Même les informations déjà géoréférencées devraient au moins faire l'objet d'un contrôle visuel. Il faudrait si possible obtenir un protocole de géoréférencement indiquant la carte de référence, l'échelle et le système de projection. Les erreurs de position sont plus facilement détectables sur les éléments linéaires (routes, rivières) que sur les polygones, qui sont fermés par plusieurs limites d'objets dépendant d'une nomenclature.

Erreurs introduites par la généralisation

La généralisation vise à simplifier, et donc à réduire le volume des informations spatiales et thématiques (JAAKKOLA 1998). Elle implique indubitablement l'introduction d'erreurs dans les données géographiques. La comparaison statistique des données source – non généralisées – et des données cible (généralisées) pour souligner le degré de généralisation n'est pas suffisante.

L'évaluation des erreurs de généralisation doit comprendre l'analyse des flux de « surfaces » d'un type de couverture à l'autre (sur le principe de l'analyse des flux de votes d'un parti vers un autre lors d'une élection) et la disparition d'éléments comme les îles ou les lacs.

Problèmes techniques

Différents environnements de données source

Les conversions répétées de toutes les données, sur la base d'un tramage par exemple, n'offrent pas la même précision géométrique et thématique si l'échelle géométrique n'est pas correctement respectée. De même, la vectorisation des données tramées ne produit pas tout à fait la même représentation que les données source si les paramètres ne sont pas vérifiés avec soin.

Ré-échantillonnage avec la même taille de pixel

Le ré-échantillonnage des données avec la même taille de quadrillage de tramage risque de mélanger des données à grande et à petite échelles. Il facilite certes le traitement, mais il ne faut pas oublier la précision et les détails « réels » (et de les décrire dans les métadonnées) lors de l'utilisation. Par exemple, le ré-échantillonnage au format 25 m x 25 m d'images satellitaires dont le tramage initial (taille des pixels) est de 30 m x 30 m (imagerie Landsat TM) n'augmente ni la précision ni l'exactitude.

Numérisation des données analogiques

Il existe plusieurs méthodes pour convertir des cartes analogiques (papier) au format numérique :

- Le balayage : il prend en compte des paramètres spécifiques (par exemple, une résolution de balayage égale à 300 dpi (points par pouce) et une résolution des couleurs de 30 bits). Les cartes tramées doivent faire l'objet d'un traitement supplémentaire pour être intégrées aux bases de données du SIG : géoréférencement, extraction des variables linéaires ou spatiales, puis vectorisation, etc.

- La numérisation avec un numériseur : la numérisation et le géoréférencement des objets géographiques à partir de cartes prennent en compte des paramètres spécifiques (par exemple, une résolution de 2 540 lpi (lignes par pouce), une précision de 0,15 mm, etc.). En fonction du logiciel de numérisation, des attributs peuvent être ajoutés directement.

- La numérisation à l'écran : la technique est comparable à celle employée avec le numériseur, mais la précision est supérieure.

Concernant la délimitation minimale lisible d'une carte source analogique à une échelle spécifique, les paramètres de numérisation ou de balayage appliqués doivent être pris en compte dans le processus d'intégration des données.

Rectification des données dans la même projection géographique

Des procédures distinctes sont utilisées pour la rectification (données vectorielles rectifiées dans un format vectoriel, données tramées rectifiées dans un format de tramage). La vectorisation offre une précision infinie au niveau ponctuel (ou linéaire), alors que le tramage couvre une certaine superficie au sol. Des problèmes cachés de conversion des modèles de données subsistent dans le processus d'intégration.

7. QUALITE

7. QUALITE

En règle générale, la qualité d'un produit se mesure à son degré de pertinence par rapport à l'objectif poursuivi. La norme ISO 8402 « Management de la qualité et assurance de la qualité – Vocabulaire », définit la qualité comme étant l'ensemble des caractéristiques d'une entité qui ont trait à sa capacité à satisfaire des besoins formulés et implicites.

Sur le plan de l'occupation et de l'utilisation des sols, cela signifie que les informations sont :

- pertinentes et utiles pour l'application ;
- fiables quant à leur contenu thématique et à leur exactitude géométrique, notamment dans le cadre d'une approche cartographique.

Le premier aspect implique une description exacte des données et la possibilité pour l'utilisateur d'accéder à des métadonnées exhaustives. Le second aspect repose sur l'évaluation de l'exactitude, en termes thématique et géométrique.

7.1. Métadonnées

Les métadonnées sont des informations sur les données, c'est-à-dire une description détaillée des spécifications. Ces « informations sur les informations » permettent d'évaluer le degré d'utilité des données pour une application spécifique.

Les métadonnées relatives à l'occupation et à l'utilisation des sols doivent contenir les informations suivantes :

- Identification des ensembles de données (titre, organisation responsable), échelle, unité de restitution
- Présentation (description générale du contexte, applications possibles, limites, etc.)
- Spécifications (couverture, nomenclature, principes de classification, méthodes et techniques de collecte, référence spatiale, unité de restitution ou cartographique, échelle, données source, périodicité)
- Paramètres relatifs à la qualité des données (procédures d'assurance qualité et mesures de contrôle)
- Politique de diffusion (prix, droit d'auteur)
- Référence des métadonnées (source des métadonnées)

Le Comité européen de normalisation a entrepris de développer une norme de métadonnées pour les informations géographiques (CEN TC 287). L'organisme MEGRIN (*Multipurpose European Ground Related Information Network*) a élaboré un répertoire descriptif des données géographiques (*Geographical Data Description Directory*, GDDD) et mis en œuvre une base de données accessible sur Internet (http://www.megrin.org/ GDDD/Overview.html). Eurostat a mis en place un dictionnaire de données pour son système d'information géographique GISCO.

7.2. Évaluation de l'exactitude

Des mesures doivent être prises pour permettre l'évaluation de la qualité des données. Ces tests produisent des relevés quantitatifs sur l'exactitude géométrique et thématique. Il existe une méthode éprouvée d'évaluation de la qualité : la validation d'un échantillon de données à partir des informations provenant d'une autre source.

La méthode d'évaluation de l'exactitude du positionnement des données géoréférencées se présente comme suit : des points de contrôle sont sélectionnés sur la carte (ou l'image) et leurs coordonnées sont comparées à celles des mêmes points tirés d'une autre source d'information (= points de contrôle au sol), tels qu'une carte topographique ou une mesure GPS sur le terrain. L'écart de position entre les points de contrôle sur l'image et sur le terrain est ensuite calculé et constitue l'erreur quadratique moyenne (EGM). À titre d'exemple, le tableau ci-dessous présente la valeur quadratique moyenne de la rectification des données satellitaires avec une résolution spatiale de 10 m et une échelle au 1/50 000e (cartographie de l'utilisation des sols dans le couloir rhodanien - projet pilote Eurostat).

	Carte X	Carte Y	Image X	Image Y	Déviation X	Déviation Y
FRANCE						
1	953675	3175750	959,00	12510,50	- 1,17	- 0,2
2	966050	3224475	2182,00	7633,50	0,58	- 0,08
3	962250	3159000	1821,25	14184,12	2,31	2,06
4	997225	3167525	5329,12	13329,62	- 0,02	- 0,03
5	970250	3162875	2625,12	13797,38	- 0,35	0,26
6	972900	3148625	2895,75	15224,50	- 0,26	- 0,64
7	1024675	3212325	8064,75	8841,62	0,88	1,23
8	973100	3199925	2897,50	10090,25	0,45	- 0,95
…	…	…	…	…	…	…
…	…	…	…	…	…	…
32	1415300	4906975	10239,50	6197,00	- 0,89	1,10
33	1387725	4849825	7969,50	12133,50	0,63	- 0,37
Erreur quadratique moyenne					**0,90**	**0,96**

Tableau 7.1 : Exemple d'évaluation de la qualité géométrique : erreur quadratique moyenne de différents points de contrôle au sol (GÉOIMAGE 1995)

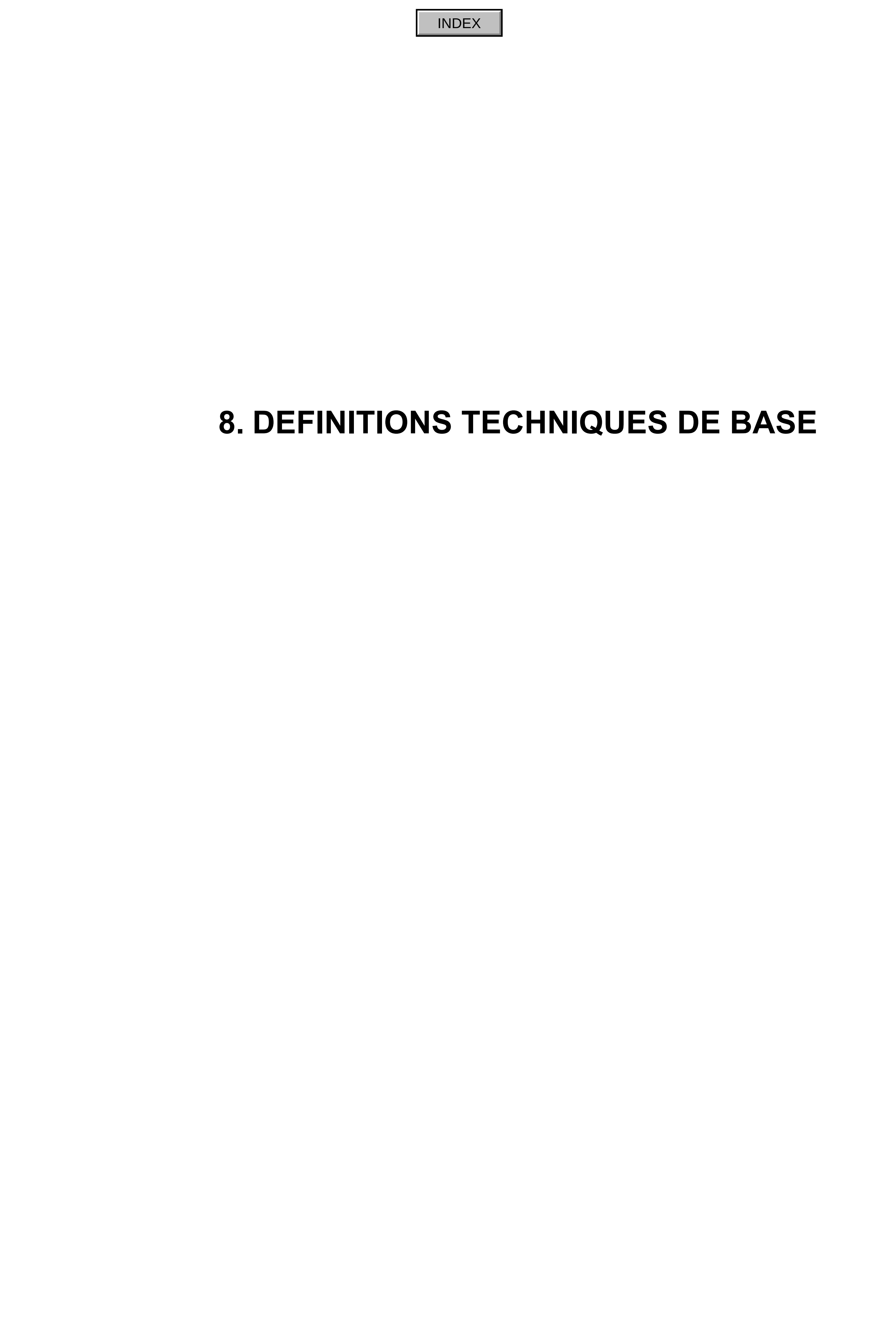
Vous pouvez également évaluer l'exactitude thématique en sélectionnant des points de test (ou pixels) et en comparant l'attribut d'occupation ou d'utilisation des sols avec la même information tirée d'une enquête sur le terrain, ou par photo-interprétation d'autres sources disponibles (clichés aériens, cartes existantes, etc.). La matrice de confusion obtenue identifie les erreurs par commission et omission, ainsi que l'exactitude des données au niveau fournisseur et utilisateur, pour chaque classe. Le tableau suivant en présente un exemple fictif.

	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6	Type 7	Type 8	Type 9	Type 10	Type 11	Total	Erreurs par commission	Exactitude utilisateur %
Type 1	53				1		1					55	2	96,3
Type 2		30			1					1		32	2	93,7
Type 3		2	12	1								15	3	80
Type 4				10								10	0	100
Type 5					54	1						55	1	98,1
Type 6					1	27	1					29	2	93,1
Type 7						4	46					50	4	92
Type 8		2						16		1		19	3	84,2
Type 9		1				2	2		15			20	5	75
Type 10							2			35		37	2	94,6
Type 11	2				1					1	20	24	4	93,3
Total	55	35	12	11	58	34	52	16	15	38	20	346		
Erreurs par omission	2	5	0	1	4	7	6	0	0	3	0			
Exactitude fournisseur %	96,3	85,7	100	90,9	93,1	79,4	88,4	100	100	92,1	100			

Tableau 7.2 : Matrice de confusion – Types d'occupation des sols (exemple fictif)

L'exactitude fournisseur (%) correspond au nombre de points **correctement classifiés** sur l'image ou la carte par rapport au **nombre total de points de contrôle au sol** pour le type d'occupation en question. L'exactitude utilisateur (%) correspond au nombre de points correctement classifiés par rapport au **nombre total de points classifiés** pour un type d'occupation spécifique ; elle prend également en compte les erreurs de classification sur la carte afin de permettre une meilleure évaluation de la qualité de la classification.

L'exactitude globale peut être évaluée, d'une part, à partir d'une moyenne des exactitudes fournisseur et utilisateur, et d'autre part, à partir du coefficient de Kappa, une statistique qui mesure l'accord, hors probabilité, entre deux cartes (le résultat d'une classification et la carte de base, par exemple) [Centre canadien de télédétection - Glossaire].



8. DEFINITIONS TECHNIQUES DE BASE

8. DEFINITIONS TECHNIQUES DE BASE

8.1. Cartes

Une carte est une représentation graphique bidimensionnelle des caractéristiques physiques (naturelles et/ou artificielles ou les deux) de tout ou partie de la surface terrestre, au moyen de signes et de symboles ou de photographies, à une échelle et dans un système de projection définis, avec indication du mode d'orientation. (DRURY 1998).

Les définitions des cartes (géographiques) incluent presque toujours le terme de représentation :

1.a : représentation sur un plan de tout ou partie d'une superficie,

1.b : représentation de tout ou partie de la sphère céleste,

2 : représentation qui évoque clairement une carte...

8.1.1. Catégorisation des cartes

Les cartes peuvent être répertoriées selon leurs différents aspects (HAKE 1982) :

- Contenu
- Origine
- Processus d'élaboration
- Échelle

8.1.2. Catégorisation des cartes selon leur contenu

Le contenu constitue le critère le plus courant de catégorisation des cartes.

Les cartes topographiques décrivent avec précision et exactitude la position géométrique des objets géographiques, tels que les zones habitées, les infrastructures (routes, voies ferrées, cours d'eau, lignes électriques, etc.), les plans d'eau, quelques aspects de l'occupation/utilisation des sols (surfaces boisées, vignobles, prairies, vergers), certains découpages administratifs, la toponymie, etc. Les cartes topographiques servent généralement de support aux cartes thématiques.

Les cartes thématiques s'appuient sur les cartes topographiques, utilisées comme support géographique, pour présenter des informations sur des sujets spécifiques ; par exemple, les cartes d'occupation des sols (basées sur les cartes topographiques), les plans directeurs d'aménagement du territoire (basés sur le cadastre), les cartes statistiques (basées sur les divisions administratives), sur la densité de population par exemple, etc.

8.1.3. Catégorisation des cartes selon leur origine

On distingue les cartes produites par l'**organisme officiel** responsable de l'étude d'un pays (cartes du « Ordnance Survey » au Royaume-Uni, du « Vermessungsämter » en Allemagne, de l'Institut Geographique National (IGN) en France ou en Espagne, de l'Institut géographique militaire en Italie, etc.), comme les cartes topographiques, des cartes réalisées par des **organismes privés** ou des sociétés, comme les cartes touristiques. Les cartes officielles servent de référence et bénéficient d'un certain caractère légal (par exemple, elles déterminent le tracé du territoire national utilisé pour établir les plans d'aménagement du territoire, les contours des sites naturels protégés, etc.).

8.1.4. Catégorisation des cartes selon leur processus d'élaboration

On distingue les « cartes de base » et les « cartes dérivées ». Les premières sont généralement des cartes à grande échelle produites par (ou pour le compte) des organismes d'enquête officiels dans le cadre de **campagnes sur le terrain** (cadastres, cartes topographiques de base, par exemple) ; les secondes sont des cartes **généralisées** à plus petite échelle, s'appuyant sur des cartes de base à grande échelle et répondant à des critères prédéfinis (cartes topographiques d'échelle moyenne, plans touristiques de ville). Dans l'idéal, les cartes de base sont des cadastres à très grande échelle (1/1 000e, par exemple) dont sont tirés les plans de ville, ou des cartes topographiques à grande échelle (1/5 000e) dont sont tirées d'autres cartes à plus petite échelle. Les cartes topographiques de base à moyenne échelle (1/10 000e - 1/300 000e) peuvent être réalisées à partir de clichés aériens grâce à des techniques de photogrammétrie et de photo-interprétation.

La *généralisation* est « la réduction des détails ou la simplification de la réalité ». Il s'agit d'une technique de traitement cartographique destinée à réduire le volume d'information présentée sur une carte, en fonction des contraintes physiques imposées par le format du papier et de l'échelle choisie.

Le même concept est appliqué aux cartes thématiques. La cartographie d'une population, par exemple, peut être réalisée à partir d'une carte topographique de base ou dérivée. Une carte généralisée de la végétation est une carte **dérivée** ; voir les exemples des figures 8.6 *(Carte de la population de Hanovre)*, 8.8 *(Carte de la population niveau NUTS)* et 8.7 *(Carte de la végétation africaine)*.

8.1.5. Catégorisation des cartes selon leur échelle

Les cartes peuvent être classées selon leur échelle.

Les principales catégories d'échelle* sont les suivantes :

Grande échelle = > 1/10 000e

Moyenne échelle = 1/10 000e au 1/300 000e

Petite échelle = < 1/300 000e

* Il n'y a en fait aucune limite fixée.

8.1.6. Exemples

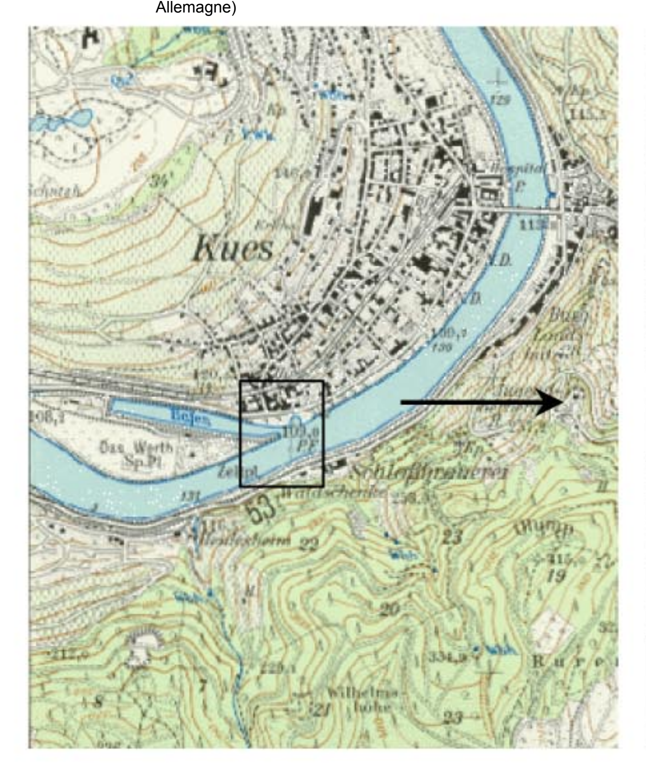
Figure 8.1: Carte topographique au 1/200 000e Feuille CC 6302 Trèves, Allemagne (HAKE 1982)

Contenu	*Origine*	*Processus de création*	*Échelle*
Carte topographique	Carte officielle (Landesvermessungsamt Rhénanie -Palatinat, Allemagne)	Carte dérivée	Moyenne, 1/200 000e



Figure 8.3 : Carte topographique au 1/25 000e Feuille 6008 Bernkastel-Kues, Allemagne (HAKE 1982)

Content	*Origine*	*Processus de création*	*Échelle*
Carte topographique	Carte officielle (Landesvermessungsamt Rhénanie-Palatinat, Allemagne)	Carte dérivée	Moyenne, 1/25 000e



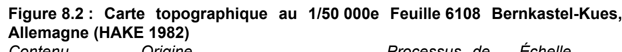


Figure 8.2 : Carte topographique au 1/50 000e Feuille 6108 Bernkastel-Kues, Allemagne (HAKE 1982)

Contenu	*Origine*	*Processus de création*	*Échelle*
Carte topographique	Carte officielle (Landesvermessungsamt Rhénanie-Palatinat, Allemagne)	Carte dérivée	Moyenne, 1/50 000e

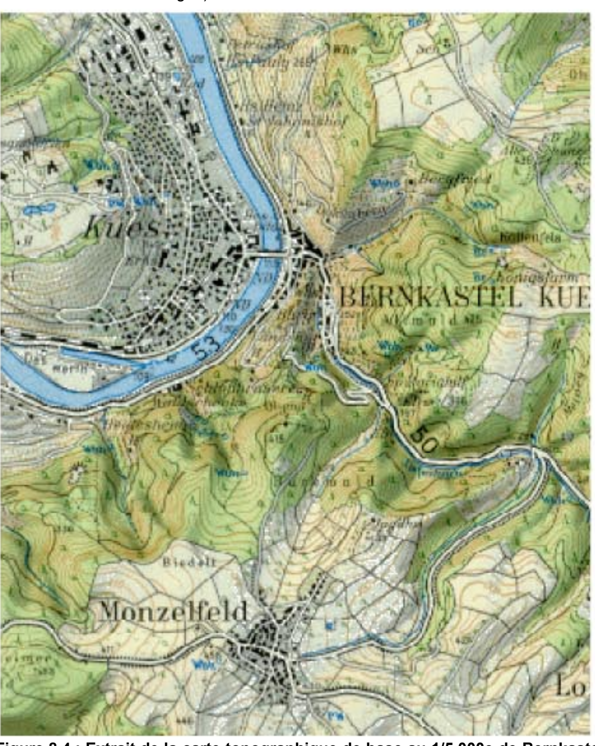


Figure 8.4 : Extrait de la carte topographique de base au 1/5 000e de Bernkastel-Kues, Allemagne (HAKE 1982)

Contenu	*Origine*	*Processus de création*	*Échelle*
Carte topographique	Carte officielle (Landesvermessungsamt Rhénanie-Palatinat, Allemagne)	Carte de base	Grande, 1/5 000e



Figure 8.5 : Carte thématique – Plan d'utilisation des sols au 1/1 000e, maison de retraite, Hambourg, Allemagne (HAKE 1982)

Contenu	*Origine*	*Processus de création*	*Échelle*
Carte thématique (Prévisions d'utilisation des sols)	Carte officielle (administration chargée de la construction et services de géodésie de Hambourg, Allemagne)	Carte de base	Grande, 1/1 000e

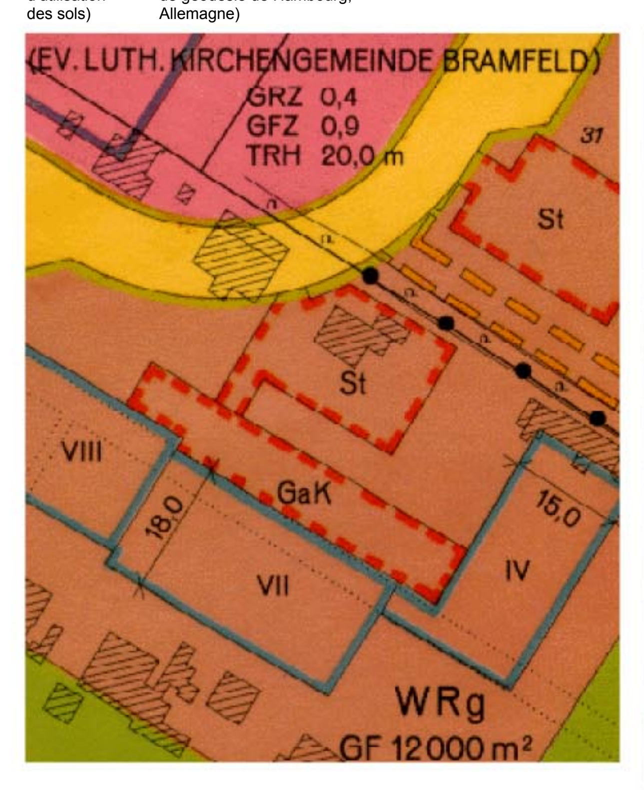


Figure 8.6 : Carte thématique – Carte de la population au 1/5 000e de Hanovre, Allemagne (HAKE 1982)

Contenu	*Origine*	*Processus de création*	*Échelle*
Carte thématique (population)	Carte officielle (Office statistique et services de géodésie de Hanovre)	Carte de base	Grande, 1/5 000e



Figure 8.7 : Carte thématique - Végétation et climat en Afrique occidentale, échelle 1/60 000 000e (HAKE 1982)

Contenu	*Origine*	*Processus de création*	*Échelle*
Carte thématique (Végétation et climat)	Carte privée (Westermann Verlag)	Carte dérivée	Petite, 1/60 000 000e

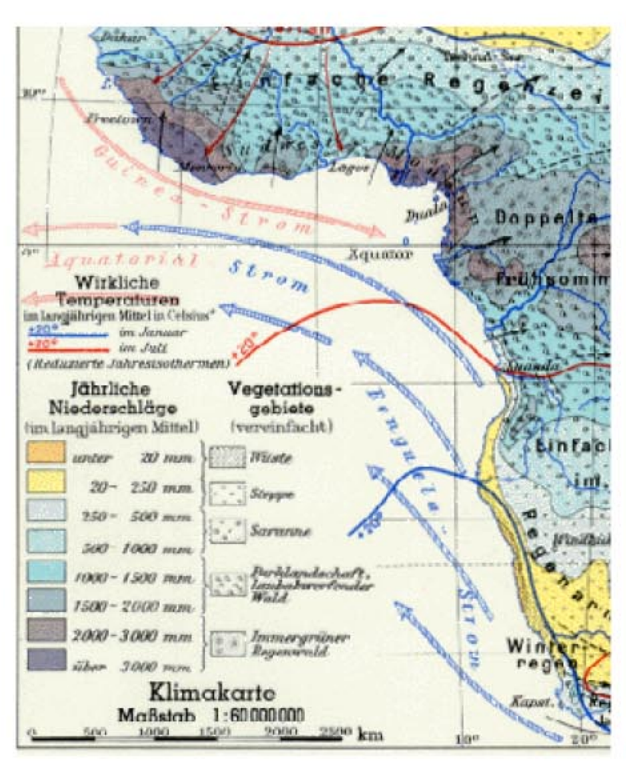
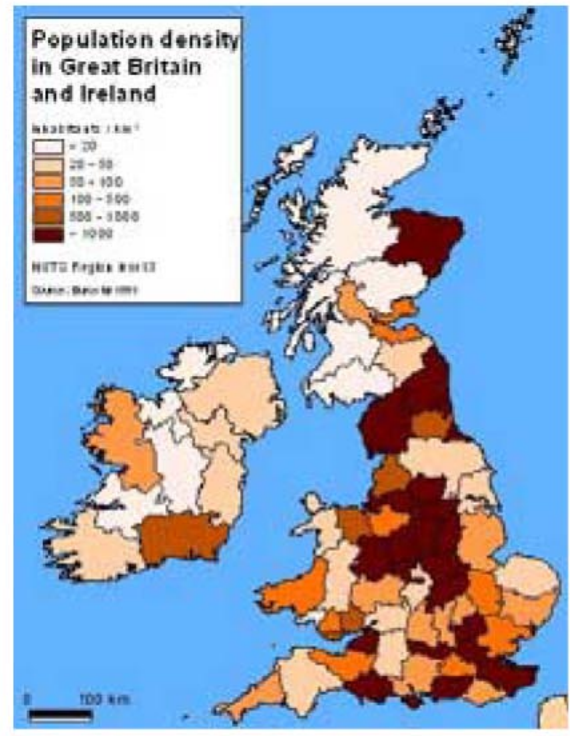


Figure 8.8 : Carte thématique – Densité de population en Grande-Bretagne et en Irlande en 1991, à l'échelle d'environ 1/25 000 000e

Contenu	*Origine*	*Processus de création*	*Échelle*
Carte thématique (densité de population)	Carte privée (CESD Communautaire, données source : CEC-Eurostat GISCO)	Carte dérivée	Petite, 1/10 000 000e



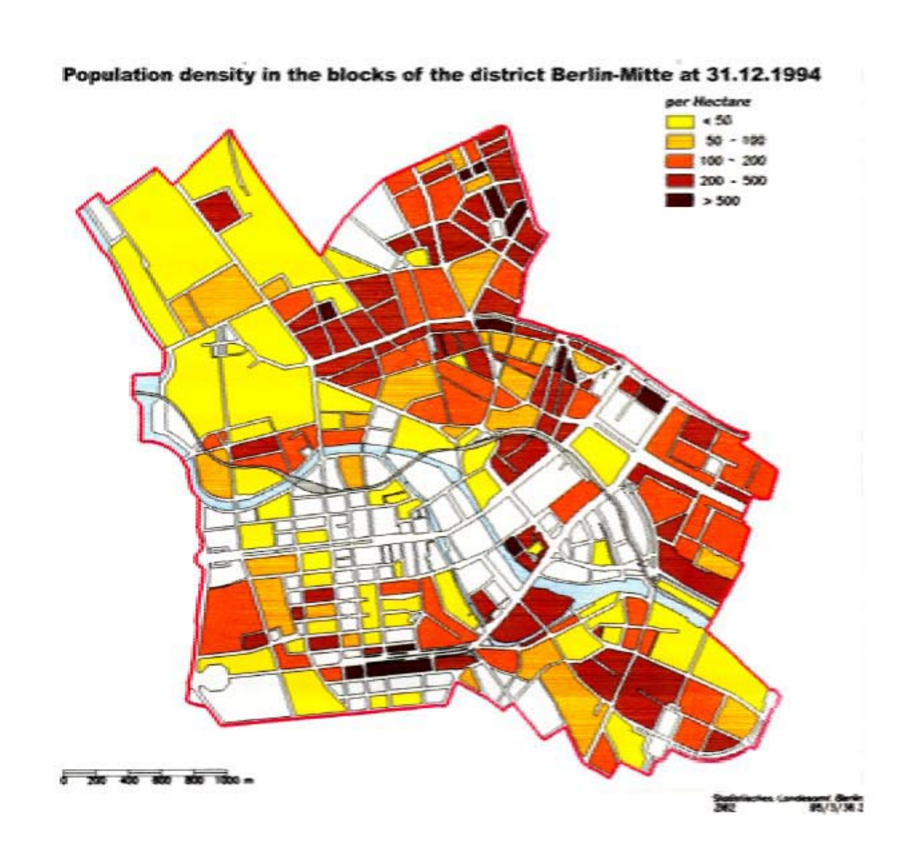
8.2. Cartes et statistiques

En ce qui concerne l'utilisation des cartes en statistique, il convient d'établir une distinction entre i) la cartographie statistique à des fins de présentation (= cartes thématiques) et ii) l'utilisation de cartes analogiques et/ou de données géographiques numériques comme élément d'un processus de collecte de données statistiques (source d'information, schéma de stratification, échantillonnage ou préparation de recensement, par exemple).

8.2.1. Cartographie statistique

La cartographie thématique est la simple utilisation de cartes pour présenter des statistiques ou des résultats d'analyse statistique. Selon le sujet, la précision géométrique et l'apparence des éléments géographiques ont plus ou moins d'importance. L'échelle requise dépend de la portée géographique et du système de référence régionale des statistiques en question. Si la superficie est le territoire de l'Union européenne et le système de référence, NUTS (Nomenclature des unités territoriales Statistiques) au niveau II, l'échelle utilisée pour présenter les statistiques (densité de population, par exemple) est différente de celle appliquée aux statistiques d'une agglomération (comme Berlin) avec le système de blocs statistiques (districts de recensement). Dans la figure 8.8 ci-dessous, les données géographiques requises sont les limites des régions NUTS, y compris le littoral, à très petite échelle (1/30 000 000e, par exemple). Pour une carte sur Berlin avec le même contenu – statistiques sur la densité de population (voir la figure 8.9), et reposant sur le système des blocs statistiques, l'échelle est au 1/5 000e environ et les données géographiques requises sont les limites des blocs.

Figure 8.9 : Densité de population dans le district « Berlin-Mitte » de Berlin au niveau des blocs statistiques (source : Statistisches Landesamt Berlin)



8.2.2. Cartes ou données géographiques utilisées comme source statistique

L'utilisation de cartes ou de données géographiques numériques comme source statistique de base est un processus plus complexe. Dans ces applications, les données géographiques sont incluses dans les statistiques et ont donc une influence notable sur le résultat. La carte ou les données doivent correspondre, en termes de contenu thématique et d'échelle (c'est-à-dire de précision et d'exactitude), à la dimension géographique des statistiques requises.

Le calcul de la superficie totale du territoire, par exemple, dépend avant tout de la notion statistique de territoire, qui peut englober des plans d'eau, des surfaces sous le niveau de la mer, des zones éloignées et inaccessibles, etc. La précision et l'exactitude des données géographiques sont le deuxième facteur important dans le calcul des surfaces à partir de bases de données numériques. Si ces informations ont été recueillies lors d'une enquête de terrain, la précision et l'exactitude sont en principe suffisantes. En règle générale, l'enquête de terrain est réalisée pour les cartographies à grande échelle. Si des cartes existantes à plus petite échelle ont été numérisées, l'échelle de la source ne devrait pas être inférieure à 1/10 000e pour les mesures de superficie.

La préparation d'un recensement, pour lequel des blocs statistiques et certaines données sur l'infrastructure (réseau routier avec noms des rues et numéros des habitations, etc.) avec un niveau de détail adéquat sont requis, est un autre exemple d'utilisation des cartes comme données de base. Si la carte n'est pas à jour, l'enquête sera incomplète. Les données géographiques de base sont essentielles à la construction d'un sondage aréolaire ou d'une stratification, par exemple pour une enquête agricole, car tout le travail statistique repose sur la qualité de la carte sous-jacente.

L'exactitude des données géographiques de base est également cruciale dans les statistiques spatiales, non seulement en termes de précision spatiale mais aussi temporelle : les limites géographiques des unités statistiques doivent être celles en vigueur à la date des statistiques. À titre d'exemple, une restructuration des unités administratives peut conduire à un redécoupage (deux unités n'en formant plus qu'une ou inversement, déplacement des limites, etc.).

Si le concept statistique est défini, les chiffres relatifs à la surface en question peuvent être calculés sur la base des données administratives tirées des registres fonciers ou du cadastre, ou par les organismes d'enquête utilisant des cartes de base à grande échelle (1/1 000e). Les bureaux de statistiques s'appuient en général sur les données fournies par l'administration ou les organismes d'enquête. Les chiffres suivants illustrent le problème posé par le concept statistique de territoire en Finlande.

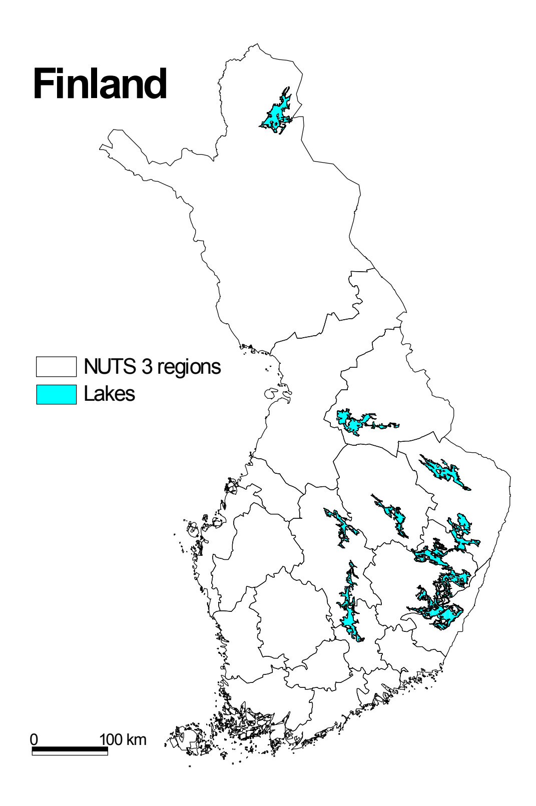


Figure 8.10 : Les lacs de Finlande[6]

Les plans d'eau représentent 33 560 km^2, soit environ 9,9 % de la superficie totale selon les statistiques de la FAO sur l'utilisation des sols.

L'utilisation de cartes dans la préparation d'une enquête statistique (recensement de la population, par exemple) réclame des conditions plus strictes, notamment des données précises et détaillées en termes de géométrie (une échelle minimale au 1/5 000e est indispensable pour pouvoir distinguer les bâtiments ; voir la figure 8.6 *Carte de la population de Hanovre*). Des renseignements détaillés sur l'usage des bâtiments seraient très utiles pour pouvoir en exclure certains (entrepôts, zones industrielles, etc.). En ce qui concerne le schéma de stratification, pour l'estimation des rendements agricoles notamment, des parcelles délimitées avec un haut degré de précision et d'exactitude géométrique sont nécessaires pour évaluer les superficies.

Dans la plupart des cas, l'analyse spatiale requiert des informations sur l'environnement, c'est-à-dire sur la topologie. Dans la mesure où les données géographiques doivent garantir un niveau de détail et d'exactitude beaucoup plus élevé, la qualité joue un rôle crucial dans la sélection des données source.

Les métadonnées, qui renseignent sur la nature de certaines variables géographiques, sont indispensables pour bien comprendre et utiliser à bon escient les informations dans des analyses et des intégrations de données spécifiques.

[6] Les données source pour cette carte sont au 1/1 000 000e. Nombre de petits lacs disparaissent à cette échelle, en raison du processus de généralisation. Les calculs effectués sur ces données géographiques de base montrent que les plans d'eau ne représentent que 3 % de la surface totale.

8.3. Système de référence géodésique et système de référence régionale

8.3.1. Système de référence géodésique

Le système de référence géodésique est défini comme un système de référence complet qui permet d'exprimer la position de tout point à la surface de la terre, y compris le système géodésique (datum), la description des coordonnées, le système de coordonnées et même une projection (CEN 1998). Toutes les cartes se fondent sur ce système de référence de positionnement : chaque point à la surface terrestre correspond à un point référencé sur la carte. En d'autres termes, le système de référence géodésique permet une description exhaustive de la surface de la terre.

La principale difficulté de cette technique cartographique consiste à représenter sur une surface plane (carte papier) l'objet tridimensionnel qu'est la Terre. Seul un modèle 3-D du globe terrestre peut éviter les distorsions des distances, des surfaces et des formes des objets géographiques. Les cartes 2-D sont toujours un compromis répondant à la finalité spécifique d'une carte.

Système géodésique

Pour pouvoir représenter la surface terrestre sur une feuille de papier, il convient de définir les paramètres relatifs à la taille et à la forme de la Terre, ainsi que l'origine et l'orientation d'un système de coordonnées. Ces paramètres constituent le système géodésique (datum géodésique).

Les descriptions des dimensions et de la forme de la Terre peuvent être de simples systèmes bidimensionnels qui prennent uniquement en compte la surface plane. Ils sont suffisants pour la cartographie à grande échelle, où les distances n'excédent pas 10 km afin d'éviter les problèmes liés à la rotondité de la Terre. L'autre extrême, à savoir la description du globe terrestre, utilise des modèles sphériques avec un rayon spécifique pour les approximations des distances à très petite échelle. Pour obtenir des mesures précises (utilisées en navigation maritime et aérienne, par exemple), il faut définir des systèmes plus complexes qui tiennent compte de la forme ellipsoïdale du globe (les pôles sont aplatis), de la topographie (le relief) et de la gravité. La figure 8.11 présente un modèle simple des dimensions terrestres.

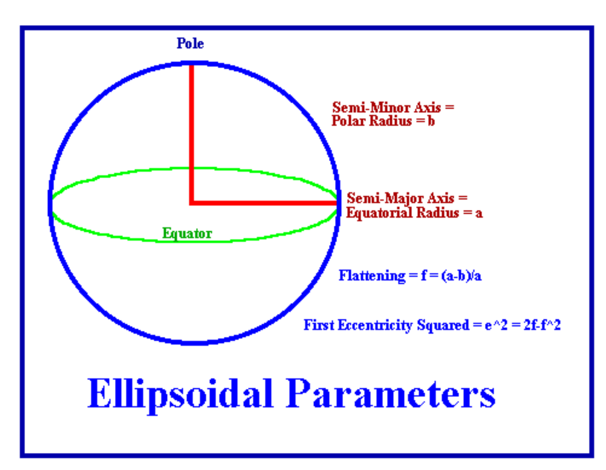


Figure 8.11 : Le système géodésique décrit la taille et la forme de la Terre. Paramètres ellipsoïdaux (adaptés de DANA 1999)

La surface topographique est très irrégulière et elle évolue dans le temps : les mouvements géologiques des plaques tectoniques sur le long terme, et les mouvements géomorphologiques à court et moyen termes, modifient le relief. Les différences de gravité et les marées modifient chaque

jour le niveau de la mer d'un lieu à l'autre, parfois de plusieurs centaines de mètres. Le géoïde - surface moyenne de la Terre – prend en compte les facteurs susmentionnés.

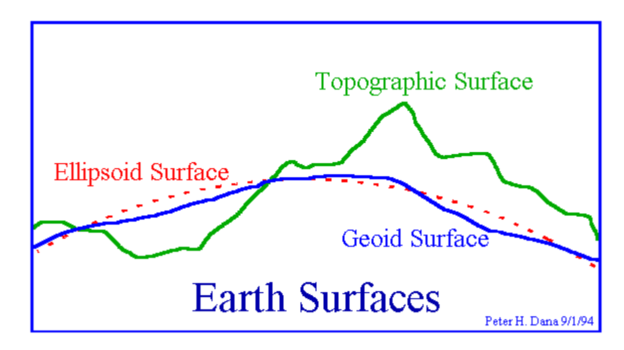


Figure 8.12 : Les surfaces ellipsoïdale, topographique et du géoïde (DANA 1999)

Comme il a été dit précédemment, des systèmes ont été définis dans différents pays pour répondre à des objectifs spécifiques. Des agences dans le monde entier utilisent divers systèmes selon l'ellipsoïde le mieux adapté à une finalité particulière et au pays concerné.

L'utilisation d'un datum inadapté au référencement des coordonnées géodésiques à partir desquelles la carte sera établie peut conduire à des erreurs de positionnement de plusieurs centaines de mètres. Le rapprochement de cartes basées sur différents systèmes se solde par des distorsions, des lacunes ou des chevauchements dans la couverture spatiale.

Les progrès technologiques réalisés dans le calcul des positionnements terrestres, avec une précision inférieure au mètre, impliquent de choisir judicieusement le datum géodésique et de convertir avec soin les coordonnées issues des différents systèmes.

Un système géodésique unifié à l'échelle mondiale est indispensable pour disposer d'informations précises au niveau intercontinental. Dans les années 1950, un programme lancé aux États-Unis a conduit à la création du système géodésique mondial « World Geodetic System 1960 » (WGS1960), développé par le département américain de la Défense. Le système WGS 72 signe une avancée majeure, après plus trois ans d'efforts considérables pour collecter des données satellitaires, des informations sur la pesanteur et des données astrogéodésiques rassemblées tout au long de l'année 1972. Il a cédé la place à WGS 84, qui utilise de nouveaux instruments plus précis et un réseau de contrôle plus étendu de stations au sol. Le WGS 84 est géocentrique : son origine est le centre massique de la Terre, océans et atmosphère inclus. Son échelle est celle du cadre terrestre local, selon la théorie relativiste de la pesanteur, et son orientation (1984.0) a été déterminée à l'origine par le Bureau international de l'heure (BIH). L'évolution temporelle de son orientation ne créera aucune rotation résiduelle au regard de la croûte terrestre (McCARTHY, D, 1996).

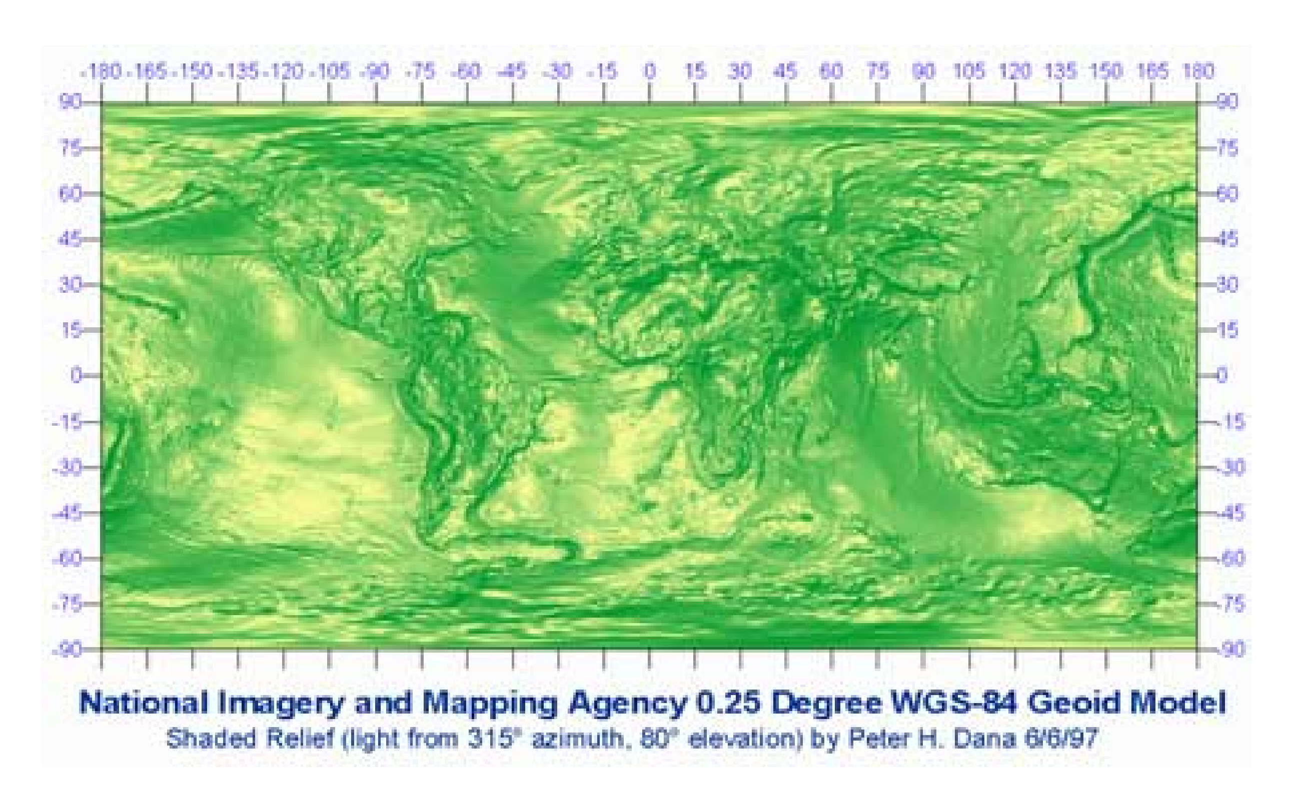


Figure 8.13 : Relief en grisé du modèle du géoïde WGS-84 (DANA 1999)

Aujourd'hui, le système géodésique mondial 1984 sert de base au GPS, le système de localisation mondial par satellite conçu par les militaires américains. Pour une description complète du WGS-84, se reporter à « National Imagery and Mapping Agency (2000): Technical Report 8350.2 3rd edition: World Geodetic System 1984.3 ».

Projections cartographiques et systèmes de coordonnées

Dans le datum géodésique, les points d'ancrage déterminent le positionnement absolu des lieux sur une carte.

L'étape suivante consiste à définir un système pour identifier un point de la surface terrestre sur une carte à partir de ses coordonnées. Les systèmes de coordonnées peuvent se fonder sur différentes **projections cartographiques**. Les points situés à la surface de la Terre (parallèles et méridiens) sont projetés sur un plan. Imaginez une sphère illuminée sur laquelle sont tracés des parallèles et des méridiens. La source lumineuse peut être interne (voir la figure 8.14) (projection gnomonique) ou située sur un point de la sphère (projection stéréographique) ou très éloignée de cette dernière (distance de la Terre au Soleil, par exemple) et émettant un rayonnement parallèle (projection orthographique).

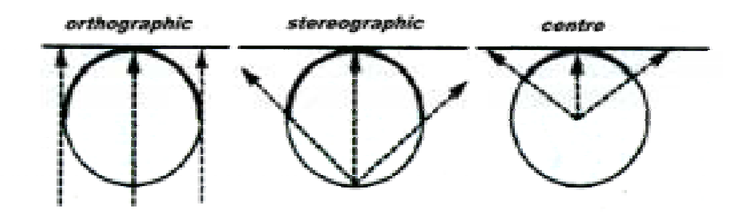


Figure 8.14 : Projections orthographique, stéréographique et gnomonique (adaptées de BRUNET 1987)

Les parallèles et les méridiens peuvent être projetés directement sur un plan – la carte papier, auquel cas la projection est dite azimutale. Le plan peut toucher la sphère (projection tangente) en un point, tel que le pôle Nord, ou la traverser (projection sécante). En roulant le papier autour du globe, on obtient une projection cylindrique ou conique. Les possibilités sont nombreuses (voir la figure 8.15).

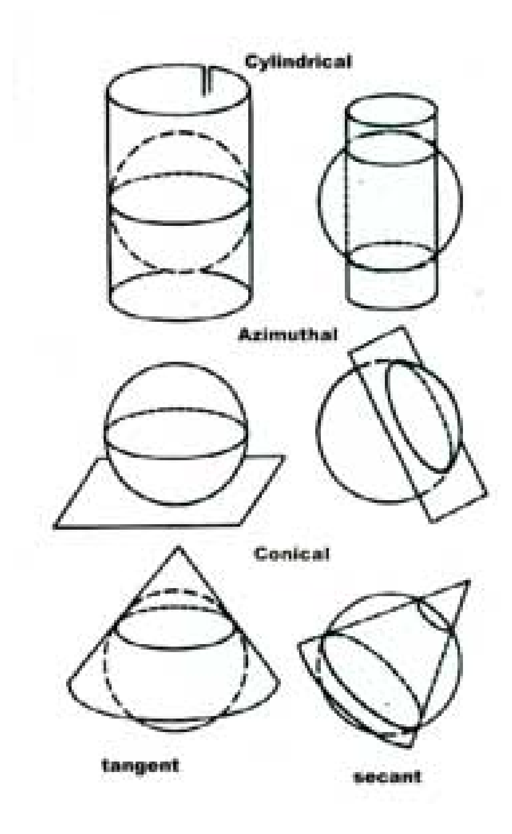


Figure 8.15 : Principes de la projection (BRUNET 1987)

Le système de quadrillage (parallèles et méridiens) projeté présentera des distorsions géométriques. Selon le type de projection et les paramètres appliqués, une ligne droite (comme une route) n'apparaîtra pas droite sur le plan et sa longueur sera inexacte ; un cercle peut se transformer en ellipse et un carré en rectangle, etc. Là encore, selon le système de projection et les paramètres mis en œuvre, l'une des conditions suivantes doit être remplie :

Équidistance : mesure des distances réelles sur la totalité de la carte. Cette caractéristique est indispensable en navigation.

Équivalence : les calculs des superficies réelles sont possibles sur la totalité de la carte. Cette propriété est importante dans la plupart des cartographies statistiques pour comparer les données des unités statistiques territoriales.

Conformalité : préservation des formes (et donc des angles) des zones géographiques.

En règle générale, le choix d'un système de projection pour réaliser une carte spécifique repose sur un compromis par rapport aux problèmes évoqués plus haut.

Selon le type de projection choisi, un système de coordonnées doit être élaboré pour permettre d'identifier sur une carte un point situé à la surface de la Terre. Les coordonnées géographiques les plus courantes sont la latitude et la longitude.



Figure 8. 16 : Système de coordonnées géographiques (DANA 1999)

De même qu'il existe différents systèmes de projection, les systèmes de coordonnées sont nombreux. En Allemagne, par exemple, le système de coordonnées géodésiques dit de Gauss-Krüger (noms de ses inventeurs) repose sur une projection conforme – c'est-à-dire que la forme des surfaces est stable – et sur un référencement utilisant l'ellipsoïde de Bessel (nom du datum géodésique). La projection UTM (*Universal Transverse Mercator*) est comparable ; elle est utilisée pour définir des positions horizontales en divisant la surface terrestre en fuseaux de 6° ; chaque zone de quadrillage est cartographiée au moyen de la projection Transverse Mercator, avec un méridien central passant au centre de la zone. Les chiffres des zones UTM désignent des bandes longitudinales de 6° comprises entre 80° de latitude sud et 84° de latitude nord. Les lettres des zones UTM désignent des bandes horizontales de 8°, qui s'étendent au nord et au sud de l'équateur. (DANA 1999)

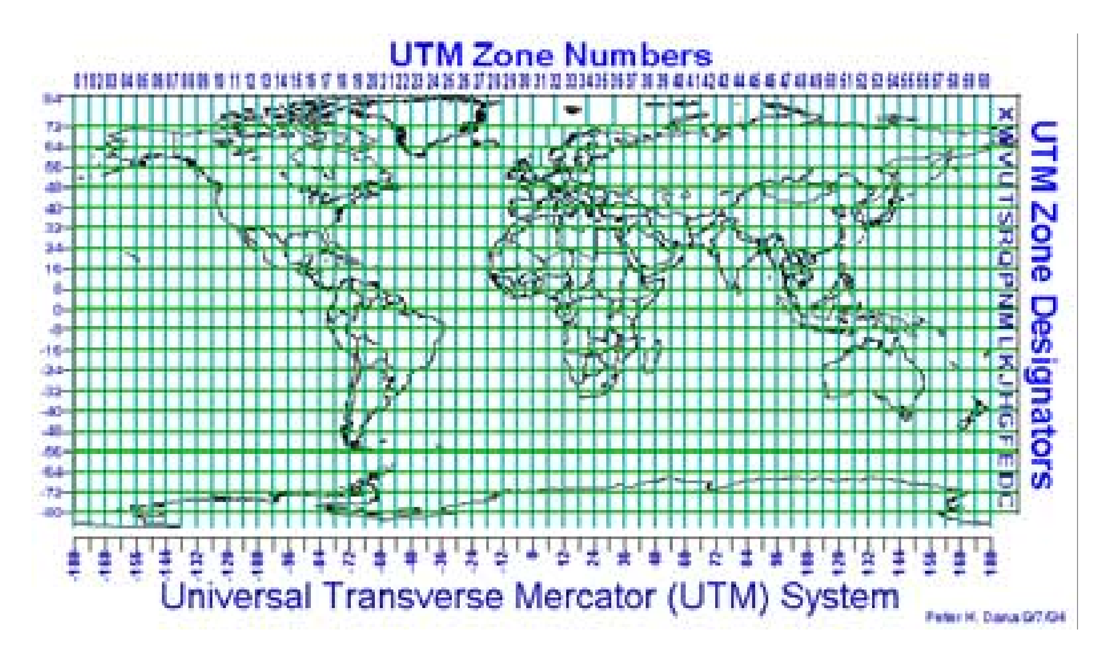


Figure 8.17 : Système de coordonnées UTM (DANA 1999)

8.3.2. Systèmes de référence régionale

En général, les données statistiques sont géocodées, c'est-à-dire associées à une surface géographique : une ville, un pays ou un groupe de pays, comme l'Union européenne. Il est prévu que les informations statistiques soient **harmonisées et comparables** sur **l'ensemble de la couverture géographique** de ce **système de référence régionale**. Dans la terminologie des géographes, le « système d'identificateurs géographiques », recoupe un concept similaire, qui définit une collection structurée d'identificateurs géographiques avec un thème et un format communs (par exemple, un code postal) (CEN 1998).

En règle générale, les statistiques font l'objet d'une **différentiation spatiale** au sein de la couverture géographique, ce qui signifie qu'elles font référence à des portions plus petites, telles que des régions statistiques, des découpages administratifs ou d'autres unités avec une fonction précise (par exemple, le trajet vers les zones de travail). Si, comme c'est souvent le cas, ces unités territoriales sont définies dans un système hiérarchique, elles peuvent être agrégées pour générer des statistiques portant sur une région plus vaste ; le système de référence régionale comprend donc plusieurs niveaux (tels que NUTS III, II et I). Plus le système est fin, plus la résolution géométrique (et le nombre d'unités) est importante et plus l'unité de restitution est petite.

Exemples de systèmes de référence régionale :



Figure 8.18 : Régions NUTS (Eurostat 1999)



Figure 8.19 : Regionales Bezugssystem Berlin, Statistisches Landesamt Berlin

L'agrégation spatiale n'implique pas nécessairement l'agrégation ou la généralisation du contenu thématique des données. À titre d'exemple, les statistiques par pays (État membre de l'Union européenne) sur l'utilisation des engrais agricoles, subdivisés en engrais azotés et engrais phosphatés, peuvent être regroupées (= additionnées) au niveau européen, avec agrégation des engrais (azotés + phosphatés) et sans agrégation des deux classes. Si un échantillon est recueilli pour toute l'Europe (niveau européen) et pour les engrais en général, il n'est pas possible de faire la distinction entre i) le niveau européen et le niveau des États membres, et entre ii) une seule classe (azotés + phosphatés) et deux classes (azotés et phosphatés).

Ceux qui ne sont pas cartographes utilisent essentiellement la notion d'échelle au sens large. Pour les statisticiens, par exemple, ce terme renvoie à la surface géographique de leur zone d'étude ou au niveau de leur système de référence régionale (tel que le niveau NUTS 2 de l'échelle européenne). Les biologistes et les écologistes peuvent penser en termes d'habitats d'espèces (une mare, une forêt ou un massif montagneux), alors que les géographes raisonneront en termes de patch dans un paysage.

La plus grande prudence est de mise afin d'éviter toute confusion lors de travaux interdisciplinaires. En cartographie, la carte à petite échelle est extrêmement généralisée en termes de géométrie **et** de contenu thématique (simplement en raison de la taille de la délimitation minimale lisible) ; ce n'est pas toujours le cas des informations statistiques en Europe, dont le contenu peut être très détaillé (en raison de l'échantillonnage) mais pas l'unité de restitution (niveau national, par exemple).

8.4. Échelle

Lors des débats interdisciplinaires entre experts, la notion d'échelle peut parfois prêter à confusion en raison des différents concepts qu'elle recouvre.

En principe, les statisticiens utilisent ce terme pour désigner le niveau du système de référence régionale. Les écologistes parlent d'« échelle spatiale » pour décrire deux caractéristiques de la collecte des données : d'une part, la **granulation** qui correspond à la résolution spatiale ou géométrique la plus fine de la collecte et, d'autre part, la **portée** qui renvoie à la taille de la zone d'étude ou à la couverture géographique du système de référence régionale du statisticien. Les adjectifs « petit » et « grand » peuvent aussi être à l'origine de malentendus. Dans l'acception courante, une application à grande échelle fait référence à un projet couvrant une vaste superficie et,

à l'inverse, une analyse à petite échelle porte sur une surface réduite. L'échelle est définie comme suit en cartographie.

L'échelle est la relation basée sur une réduction linéaire entre une carte et la réalité. Pour calculer l'éloignement entre deux points dans le monde réel, la distance mesurée sur la carte doit être multipliée par le facteur d'échelle ; par exemple, une échelle au 1/1 000 000e signifie qu'une distance de 1 cm sur la carte correspond à 1 000 000 cm = 10 000 m = 10 km dans la réalité. En l'absence d'échelle, une carte n'est plus qu'un diagramme. Comme nous l'avons déjà indiqué dans les chapitres précédents, une carte présente toujours des distorsions géométriques puisqu'elle est une représentation plane d'un monde tridimensionnel. Les distorsions peuvent affecter les mesures de distance et de superficie, ainsi que la forme des zones délimitées, selon la projection et le système de coordonnées utilisés sur la carte. Au sens étroit, l'échelle d'une carte est une approximation portant sur toute la superficie représentée puisque, pour chaque point, l'échelle réelle peut être calculée à l'aide des paramètres exacts des coordonnées, de la projection et du datum géodésique à ce point sur la surface terrestre.

L'échelle peut s'exprimer de trois manières :

1. Indication : telle que « 1 cm pour 10 km », facile à comprendre, mais nécessite une règle ;
2. Échelle de distance : exprimée en km ou miles – très utile pour la reproduction graphique puisque l'échelle change à l'impression ;

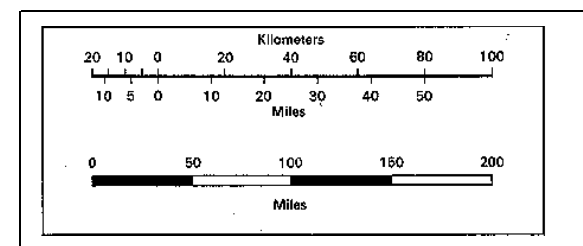
	L'extrémité gauche de l'échelle est subdivisée en unités plus petites afin de faciliter l'estimation de distances précises.

Figure 8.20 : Exemples d'échelles de distance (université du Texas, Austin - Département de géographie 1999)

3. Échelle numérique : exprime le rapport de la réduction sous la forme d'une fraction, telle que 1/100 000e ou 1/100 000e (sans indication de l'unité) ; elle sert à identifier des séries de cartes, mais peut être une source de confusion puisque la partie droite de la fraction augmente au fur et à mesure que l'échelle décroît :

   l'échelle de 1/1 000e est supérieure à 1/10 000e, à 1/100 000e, à 1/1 000 000e, etc.

Le passage d'une grande à une petite échelle implique un processus de *généralisation*, c'est-à-dire la réduction du volume d'informations présentées sur la carte du seul fait de la limitation de l'espace disponible sur le papier. Les polygones (surfaces) deviennent des points, les deux berges d'une rivière sont réduites à une seule ligne, certains détails sont supprimés.

Échelle dans les bases de données spatiales

Au sens strict, la notion d'échelle est un concept cartographique qui **ne s'applique pas** aux représentations abstraites enregistrées dans une base de données géographiques. Dans l'idéal, le contenu d'une base de données géographiques est affranchi de toute échelle, c'est-à-dire que les coordonnées d'un point (ligne, polygone) sont le reflet de la réalité. En fait, les données sont souvent numérisées à partir de cartes qui ont été établies sur la base d'une échelle spécifique, et l'erreur de géométrie se propage alors dans le système numérique.

Les concepts appliqués dans les *BASES DE DONNEES SPATIALES* sont les suivants :

- Précision : degré de détail dans la capture d'une mesure.
- Exactitude : relation entre une mesure et la réalité représentée.
- Résolution : le plus petit objet qui peut être représenté.

Dans la terminologie de la *TELEDETECTION*, la **résolution spatiale** est la plus petite surface au sol (unité territoriale) dont le signal radiométrique est détectable par le capteur. La taille de cette unité dépend des caractéristiques du capteur et de l'altitude de la plate-forme. Lors des discussions entre statisticiens et spécialistes en télédétection, ce terme est parfois confondu avec « *unité d'observation »* ou *« unité cartographique ».*

GOODCHILD, M (1991) : Issue of quality and uncertainty. In: Advances in Cartography.- London (ed. MÜLLER) pages 113-139

Le contenu d'une base de données géographiques n'a pas besoin d'être généralisé, si ce n'est à des fins de présentation (sur papier ou écran).

8.5. Unités

8.5.1 Unités des cartes géographiques

Les unités des cartes géographiques désignent des unités de mesure métrique dans lesquelles les données géographiques sont enregistrées (format analogique ou numérique) : pouces, pieds, mètres ou degrés, minutes et secondes, etc.

8.5.2 Unité cartographique

L'unité cartographique est la plus petite surface mesurée et cartographiée. Sa taille est avant tout déterminée par les objectifs de l'utilisateur, en l'occurrence l'échelle de la carte à établir, c'est-à-dire par la précision et l'exactitude requises pour identifier les plus petits objets sur la carte. Elle dépend aussi de la précision géométrique ou de la résolution spatiale autorisée par la technique ou l'outil de collecte de données, qui doit donc être choisi en conséquence.

La précision et l'exactitude des mesures terrestres peuvent être de quelques millimètres ou centimètres. C'est la cas, par exemple, pour le réseau de points trigonométriques, un schéma d'installation ou un plan de propriété. D'autres techniques, comme la photogrammétrie des clichés aériens ou des images satellitaires, fournissent des unités cartographiques inférieures à un ou à plusieurs mètres, selon la résolution spatiale. En principe, pour pouvoir différencier un élément de la surface observée sur des images de télédétection optique, celui-ci doit être de dimension égale ou supérieure à la cellule de résolution. L'attribution d'un pixel à une classe d'occupation des sols, par exemple, implique l'analyse des pixels adjacents, au minimum dans une fenêtre de 3 x 3 pixels. Dans la pratique, un filtrage est effectué suite à une procédure automatique ou semi-automatique de classification (analyse d'agrégats, classification par maximum de vraisemblance, etc.), afin d'éliminer tout « bruit » au niveau des pixels. Les récents progrès réalisés dans la classification des images de télédétection permettent d'identifier d'abord les segments de valeurs spectrales homogènes, puis de les classifier. L'unité cartographique correspondra alors à la taille minimale de ces segments.

8.5.3. Délimitation minimale lisible

La délimitation minimale lisible d'une carte est la taille minimale identifiable d'une ligne ou d'un polygone sur le papier. Une ligne noire doit avoir une épaisseur de 0,05 mm et une ligne colorée, de 0,08 à 0,1 mm. Quant aux dimensions spatiales des polygones, la largeur doit être de 0,3 mm, et de 1 mm² pour les superficies en couleur.

Épaisseur d'une ligne noire :	0,05 mm	
Épaisseur d'une ligne colorée :	0,08 à 0,1 mm	
Distance entre lignes :	0,25 mm	
Dimensions des polygones :	0,3 mm	
Espace entre polygones pleins :	0,2 mm	

Les figures sont surdimensionnées.

Selon l'échelle de la carte, cette unité représente des surfaces variables dans la réalité : au 1/1 000e, la plus petite surface représentée est égale à 1 m² (1 mm x 1 mm sur la carte) ; au 1/100 000e, la même surface correspond à 10 000 m² (= 1 ha). Par conséquent, la précision et l'exactitude des mesures terrestres en centimètres seront biaisées sur une carte au 1/100 000e : 1 mm correspond à 100 mètres dans la réalité. Le même concept s'applique aux enquêtes sur le terrain : une ligne tracée sur le papier représente une certaine surface, selon l'échelle de la carte à établir.

La notion de délimitation minimale lisible est parfois confondue avec celle d'unité cartographique.

8.5.4. Unité d'observation

Dans les enquêtes statistiques, les unités d'observation correspondent aux sujets d'étude ; par exemple, des êtres humains, des familles, des ménages, des plantes ou animaux, des stations climatiques, des parcelles agricoles, des divisions administratives ou autres unités territoriales, etc. Les caractéristiques d'investigation, les variables, sont très diverses : âge, nombre d'enfants, revenus du ménage, ampleur des dégâts (causés aux arbres, par exemple), durée d'ensoleillement, occupation ou utilisation des sols, etc.

8.5.5. Unité de restitution

En règle générale, l'unité de restitution est celle dans laquelle les statistiques sont représentées. Elle peut avoir plusieurs dimensions temporelles (intervalle de temps) et spatiales (régions).

Les données recueillies lors d'un recensement de la population d'un pays peuvent être présentées pour chaque ménage. L'unité d'observation est alors égale à l'unité de restitution, comme c'est le cas dans plusieurs pays nordiques. Les résultats peuvent aussi être regroupés dans l'espace jusqu'à un niveau supérieur du système de référence régionale (le niveau NUTS 4, par exemple, devient l'unité de restitution). Les données tirées d'un sondage à base aréolaire sont représentatives d'une certaine unité de référence régionale, en fonction du mode d'échantillonnage.

Lors de travaux interdisciplinaires, l'un des problèmes les plus épineux est la confusion ou l'erreur d'interprétation des termes, pourtant définis avec précision mais sur la base de concepts sectoriels différents. Cette situation se produit en particulier quand des termes similaires, voire identiques, sont employés.

Erreurs d'interprétation ou confusion peuvent survenir au cours de discussions entre statisticiens et géographes (cartographes, spécialistes en télédétection) sur des termes tels que « unité cartographique », « unité d'observation », « délimitation minimale lisible », « résolution spatiale » et « unité de restitution ».

Comme indiqué plus haut, l'unité cartographique désigne toujours une surface dans la terminologie de la cartographie. La taille requise dépend de l'échelle de la carte à établir (c'est-à-dire de la taille des objets à représenter), laquelle conditionne la technique de collecte des données (photogrammétrie terrestre ou aérienne). Si un très haut degré de précision ou d'exactitude géométrique est requis – pour un schéma d'installation montrant les conduites d'eau ou le réseau d'alimentation électrique, afin de permettre l'aménagement de sites de construction, par exemple – la carte dressée doit être à une très grande échelle, au 1/100e ou 1/500e. Si le thème est la propriété des parcelles de terrain, l'échelle adéquate peut être au 1/1 000e ou 1/5 000e. Les données géographiques servent uniquement de base aux cartes thématiques, par exemple celles montrant la densité de population des États membres de l'Union européenne associés aux régions NUTS **II** ; la portée géographique considérée, en l'occurrence la totalité du territoire de l'UE, et l'espace disponible pour imprimer la carte, déterminent l'échelle appliquée (par exemple 1/30 000 000e pour un format A4). Dans ce cas, la précision et l'exactitude géométriques ne sont pas un facteur important.

Si, pour cette même carte, l'unité de restitution est le district de recensement (soit le niveau le plus fin du système de référence régionale), la délimitation minimale lisible jouera un rôle dans la détermination de l'échelle : un district de 100 ha sera représenté sur une carte au 1/10 000 000e par une superficie de 1 mm², soit la délimitation minimale lisible d'un point coloré sur le papier.

Comme indiqué plus haut, les unités d'observation des études statistiques **peuvent être** des surfaces délimitées (parcelles, habitations, agglomérations urbaines, divisions administratives comme des communes, des régions NUTS ou des pays). Il convient d'opérer une distinction entre la référence régionale d'une unité d'observation et une unité d'observation qui **est** une région.

En terminologie statistique, l'unité d'observation comprend presque toujours une référence régionale : les statistiques ont en principe une portée géographique, que ce soit la totalité du territoire d'un pays (population française) ou une zone plus petite délimitée avec précision (chômage à Dublin). Elles peuvent aussi avoir une référence géographique plus fine. Un système de référence régionale permet de comparer les propriétés des unités d'observation dans différentes régions. Reposant généralement sur une organisation hiérarchique, il se compose d'adresses, au plus petit niveau de référence, de blocs statistiques (ou districts de recensement), de districts, communes, départements et provinces, jusqu'au niveau national et international (UE, OCDE, États membres de l'OMC, etc.). Il est bien évidemment indispensable que les statistiques soient harmonisées, voire standardisées sur l'ensemble de la couverture géographique avant toute comparaison.

8.6. Statistiques géocodées et informations géographiques

En principe, toutes les statistiques sont *géocodées*, c'est-à-dire qu'elles se rapportent à un système de référence régionale utilisant des codes alphanumériques pour répertorier les surfaces ou les lieux (noms, numéros, etc.) : les données nationales font référence à l'ensemble d'un pays, les données régionales à des surfaces moins vastes (essentiellement des divisions administratives, comme les trois niveaux des régions européennes NUTS).

Les données géographiques fournissent des informations sur :

- des lieux situés à la surface de la Terre,
- la position d'un objet,
- l'objet situé à un emplacement donné.

(GOODCHILD 1997)

Les données géographiques sont représentées sur des cartes papier au format analogique et sont soumises à différentes contraintes de présentation et de contenu de l'information, selon l'échelle, le degré de généralisation et le type de carte. À l'heure actuelle, elles sont traitées dans des systèmes informatiques, à l'instar d'autres types d'information.

Dans l'idéal, les objets géographiques ont des limites spatiales et des attributs bien définis ; par exemple, des parcelles avec des délimitations précises et exactes levées sur le terrain, et différents attributs tels que la propriété, l'usage effectif, l'usage autorisé, la valeur fiscale, etc., qui s'appliquent uniformément à l'objet dans son intégralité[7].

Points, lignes et surfaces délimitent ces objets géographiques dans un système de référence absolu. Les lignes sont constituées de points, avec les coordonnées exactes du système de référence absolu ; les surfaces (polygones) sont composées de lignes. Les informations sur la délimitation spatiale de ces entités sont appelées données géographiques.

Les SIG analogiques (= cartes) présentent ces points, lignes et surfaces (= polygones) à des échelles spécifiques sur le papier. Les SIG numériques enregistrent les coordonnées réelles des objets géographiques au format numérique.

Un système d'information géographique (SIG) est :

- un système d'acquisition, de stockage, de manipulation et de diffusion des données géographiques,
- une catégorie de logiciel.

Un exemple pratique de SIG combine logiciel, matériel, données, utilisateur, etc., pour résoudre un problème, étayer une décision, contribuer à l'élaboration d'un plan.

(GOODCHILD 1997)

8.7. Géoréférencement, géocodage

8.7.1 Géoréférencement

Comme indiqué précédemment, les objets géographiques sont délimités par des points, des lignes (une suite de points reliés entre eux) et des surfaces (une suite de lignes reliées entre elles). Le *géoréférencement* désigne l'attribution des coordonnées d'un système de référence géographique absolu à ces points (lignes, polygones). Par exemple, les coordonnées géographiques de l'aéroport de Lisbonne sont (approximativement) 38°45' nord (de l'équateur) et 9°19' ouest (de Greenwich).

- En TELEDETECTION, le géoréférencement concerne le prétraitement de l'image à des fins de correction géométrique :
- La « *correction géométrique* » est le traitement qui corrige les distorsions spatiales d'une image.
- La « *superposition de zonages* » est le processus qui consiste à superposer plusieurs images ou photographies de manière à faire coïncider des points géographiques équivalents.

[7] … d'autres phénomènes géographiques sont plus souvent considérés comme des surfaces continues – pression atmosphérique, relief représenté par la surface hypsométrique, charges hydrauliques ou émissions polluantes. Ils sont généralement représentés par des surfaces mathématiques lisses (souvent des fonctions polynomiales) qui varient en continu et uniformément dans l'espace et le temps. (BURROUGH & FRANK 1996 p. 4). La délimitation de ces phénomènes est un autre problème.

- Le « *géoréférencement* » est la transformation de l'image (chaque pixel) en fonction d'un système cartographique de référence.

Ces techniques impliquent l'attribution, le calcul ou le « ré-échantillonnage » d'une nouvelle valeur pour chaque pixel déplacé sur l'image tramée : à l'aide, par exemple, de points de contrôle au sol dont les coordonnées sont connues et des points de contrôle correspondant sur l'image, il est possible de calculer les coordonnées géographiques (c'est-à-dire la localisation précise et exacte du pixel) dans une nouvelle matrice géoréférencée. Le calcul et l'attribution d'une nouvelle valeur spectrale pour la position du pixel dans l'image géoréférencée s'appelle le « *ré-échantillonnage* ». Il s'appuie sur les valeurs des pixels voisins non corrigés dans la matrice d'origine. Le ré-échantillonnage peut être réalisé simplement à partir de la valeur du pixel le plus proche ou par interpolation, en utilisant une fenêtre autour de la précédente position du pixel.

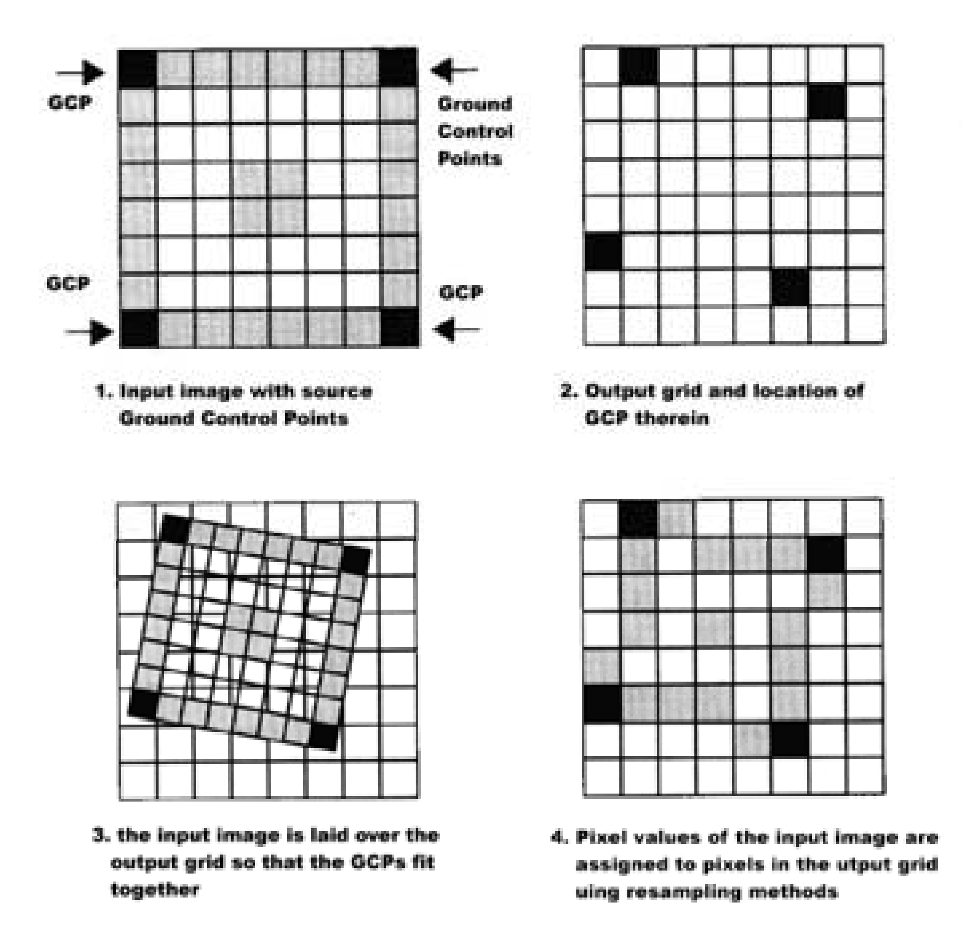


Figure 8.21 : Principe du géoréférencement (ERDAS 1994)

8.7.2. Géocodage

Le *géocodage* est l'attribution d'un code (nom ou numéro, par exemple) à une entité géographique.

En STATISTIQUE, les termes « géocodage » et « géoréférencement » sont employés dans un sens plus large : une statistique est référencée ou codée en fonction d'une référence régionale, le plus souvent une unité territoriale, comme un bloc statistique ou une région NUTS (voir le paragraphe « Système de référence régionale »). Le géocodage de l'aéroport de Lisbonne est RC132 'Grande Lisboa' au niveau NUTS 3.

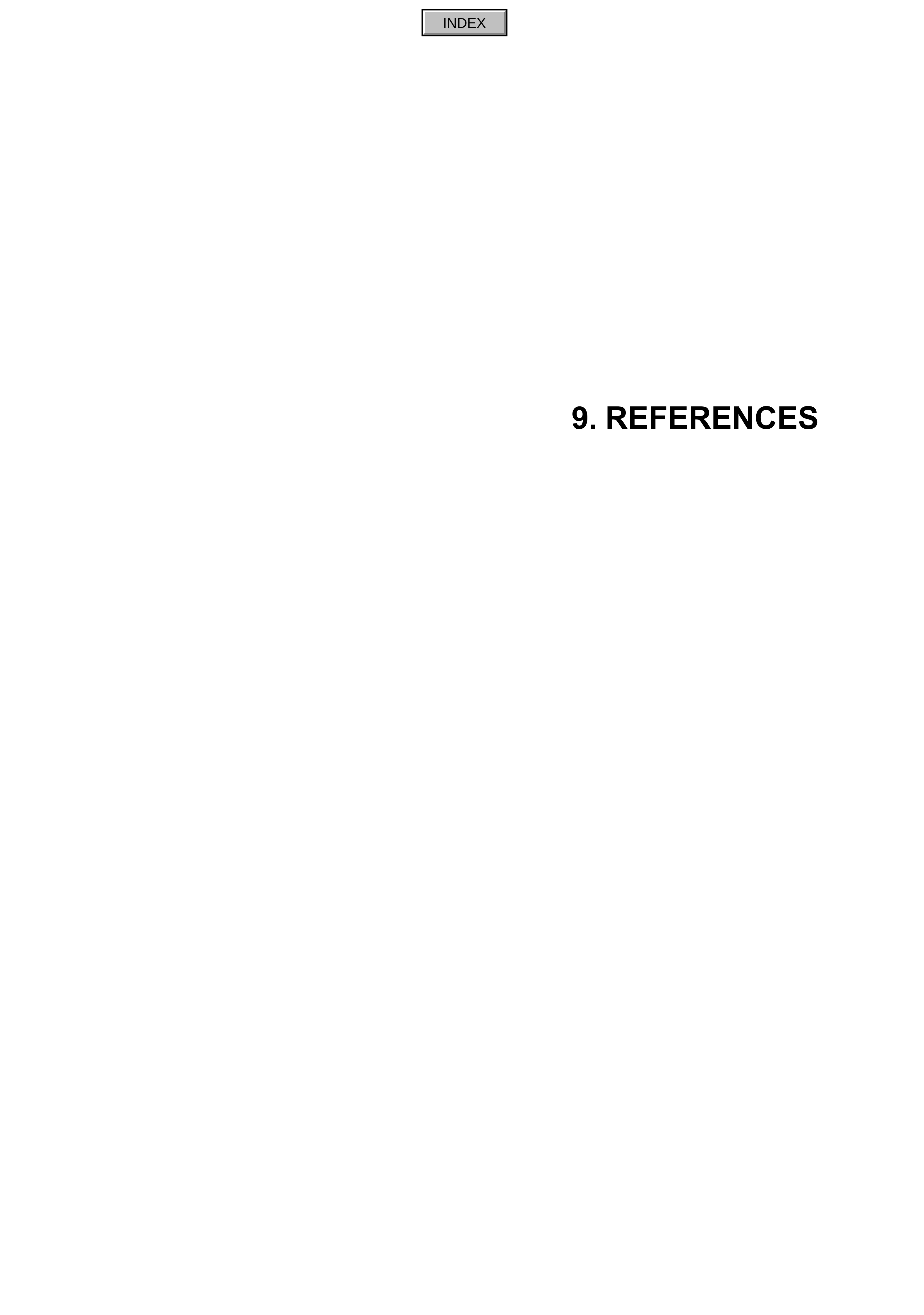
Lors d'un recensement de population, les données correspondant aux unités d'observation – les individus – peuvent être associées (géocodées) à l'adresse postale (grande précision) ou au bloc

statistique (agrégation spatiale). Il n'y a pas de géoréférencement au sens strict – attribution de coordonnées géographiques – puisque les données peuvent être enregistrées dans des tables avec les adresses ou les codes des districts de recensement. Les informations requises pour géoréférencer le pixel correspondent aux coordonnées géographiques des limites des unités de référence (degrés, minutes, secondes ou mètres utilisant un système de coordonnées défini). Le géoréférencement permet de créer des données topologiques qui reflètent les relations de voisinage entre les régions de référence ; cette approche est cruciale dans certains types d'analyse spatiale.

Dans le cas d'un sondage à base aréolaire, les informations recueillies sur une partie de la zone sont représentatives de l'ensemble et assignées (géocodées) à toute la superficie[8]. S'il est nécessaire de géocoder les statistiques avec plus de précision, à un niveau inférieur du système de référence régionale, un échantillon représentatif doit être défini selon la taille (en termes de superficie) et le nombre des unités de référence aux différents niveaux de ce système. Le mode d'échantillonnage détermine l'unité de restitution spatiale minimale.

---

[8] À titre d'exemple, la base de sondage aréolaire MARS est conçue pour permettre des estimations au niveau européen. Les échantillons ne sont pas représentatifs au niveau national (celui de chaque État membre de l'UE).



9. REFERENCES

9. REFERENCES

ABEL, D. J., BENG CHIN O., KIAN-LEE T. & SOON HUAT T. (1998) : Towards integrated geographical information processing – International Journal on Geographic Information Science, vol. 12, n° 4 1998, pages 353-371

ALBERTZ, J. (1991) : Grundlagen der Interpretation von Luft- und Satellitenbildern : eine Einführung in die Fernerkundung. Wissenschaftliche Buchgesellschaft Darmstadt

ALUN, J., CLARK, J. (1997) : Driving forces behind European land use change: an overview. CLAUDE Resource Paper n° 1

BÄHR, H.-P. (1985) : Digitale Bildverarbeitung: Anwendung in der Photogrammetrie und Fernerkundung - Karlsruhe

BAULIES, X.I., SZEJWACH, G. (ed.) (1997) : Survey of needs, gaps and priorities on data for land use and land cover change research. LUCC Data requirements workshop, Barcelone 11-14 Novembre 1997, LUCC report series 3

BRUNET, R. (1987) : La Carte Mode d'Emploi.- Paris

BUNGE, M. (1983) : Epistemology and methodology I: exploring the world, vol. 5 of Treatise on basic philosophy, Dordrecht and Boston, Reidel

BURROUGH, P.A. & FRANK, A.U. (1996) : Geographic objects with indeterminate boundaries.- GISDATA series European Science Foundation, London, Bristol

CEN (1998) : Geographic information - Vocabulary. - CEN Report CR 13436:1998, Paris

Centre canadien de télédétection (1999) : Notions fondamentales de télédétection – Cours tutoriels de télédétection

Centre canadien de télédétection (1999) : Glossaire - http://www.ccrs.nrcan.gc.ca/ccrs/eduref/ref/glosndxe.html

CHRISMAN, N R. (1998) : Rethinking levels of measurements for cartography – Cartography and Geographic Information Systems, Vol. 25, n° 4, 1998, pages 231-242

COCHRAN, W.G. (1977) : Sampling Techniques. Wiley, New York, 3rd edition

CONGALTON, R. (1991) : A review of assessing the accuracy of classifications of remotely sensed data – Remote Sensing of Environment, Vol. 37, pages 35-46

COTTER, J. ; NEALON, J. (1987) : Area Frame Design for Agricultural Surveys. United States Department of agriculture, National Agricultural Statistics Service, Research and Applications Division

COWEN, D. J. (1997) : NCGIA Core Curriculum in GIScience

CROI, W. (1999) : European Land Cover and Land Use Information Database – Working Party Land Use Statistics May 1999 (Doc. LAND/19)

DANA, H. P. (1999) : The Geographer's Craft Project, Department of Geography, université du Texas, Austin http://www.utexas.edu/depts/grg/gcraft/notes/datum/datum_ftoc.html#Contents

DI GREGORIO, A. & JANSEN, LJM. (1997) : A new concept for a Land Cover Classification System. Earth observation and evolution classification, 1997, Compte rendu de la conférence des 13-16 octobre 1997 à Alexandrie, Égypte, 10 pages

DEVOGELE, T., PARENT, C. & SPACCAPIETRA, S. (1998) : On spatial database integration - International Journal on Geographic Information Science, Vol. 12, n° 4 1998, pages 335-352

DRURY, S.A. (1990) : A Guide to Remote Sensing - Oxford, 199 pages

DUHAMEL, C. & CROI, W. (1998) : LANES - Development of a harmonised framework for multi-purpose lan cover / land use information systems derived from earth observation data – Rapport final (non publié)

DUHAMEL, C. (1998) : CESD Communautaire: Background paper on nomenclatures, LANES concerted action, Research program on the Environment and climate, Commission européenne

DUHAMEL, C. & VIDAL, C. (1998) : Objectives, tools and nomenclatures in proceedings of Eurostat seminar Land Cover and Land Use Information Systems for European Union Policy needs, Luxembourg 1998

DUPRAT, H. (1972) : Les conditions linguistiques du transfert technologique – Revue Automatisme Tome XVII, n° 11, Paris

ECO, U. (1988) : Le Signe, Éditions Labor, Bruxelles

EKHOLM, A. (1996) : A conceptual framework for classification of construction works

EUROSTAT (1992) : Task force statistiques spatiales, rapport final Statistiques d'utilisation des sols

EUROSTAT (1995) : Régions - Nomenclature des Unités territoriales statistiques NUTS - Luxembourg, non publié

EUROSTAT (1996) : CLUSTERS (Classification of Land Use STatistics - Eurostat Remote Sensing) : a pilot nomenclature – A statistical framework for remote sensing applications in statistics - non publié

EUROSTAT (1999) : Nomenclature des Unités territoriales statistiques NUTS - Luxembourg

Eurostat (1998) : The GISCO Database Manual.- non publié

Erdas (1994) : Field Guide. - Atlanta, GA, USA

FAO (1996) : Multiple frame agricultural surveys. Current surveys based on area and list frame sampling methods. FAO Statistical Development Series, Vol. 1

FAO (1998) : Multiple frame agricultural surveys. Agricultural survey programs based on area frame or dual frame (area and list) sample designs. FAO Statistical Development Series, Vol. 2

FRITSCH, D., GLEMSER, M., KLEIN, U., SESTER, M., STRUNZ, G. (1998) : Zur Integration von Unsicherheit bei Vektor- und Rasterdaten - Geographische Informationssystems 4 / 1998, pages 26-35

GALLEGO, F.J., CARFAGNA, E. (1998) : Research project on Geographic sampling strategies and remote sensing – Rapport final – non publié

GALLEGO, F.J. (1995) : Sampling Frames of Square Segments. Office des publications officielles des Communautés européennes, Luxembourg

GÉOIMAGE (1995) : Utilisation de la télédétection pour l'inventaire, le suivi et la gestion du couloir rhodanien - Rapport final – Projet Eurostat, non publié

GOODCHILD, M. F. (1997) : What is Geographic Information Science? - NCGIA Core Curriculum in GIScience, http://www.ncgia.ucsb.edu /giscc/units/u002/u002.html, sorti le 7 octobre 1997.

GOODCHILD, M. F. ; LONGLEY, P. A. ; MAGUIRE, D. J. ; RHIND, D. W. (1999) : Geographic Information Systems. - 2nd edition, New York

HAKE, G. (1982) : Kartographie - Berlin, New York

HJELMSLEV, L. (1959) : Pour une sémantique structurale - Essais linguistiques, Copenhagen, Nordisk Sproog-OG Kulturforlag

HUNTING Technical Service & The Centre for Agricultural Strategy (1997) : Quality Control Procedures for Land Use Data derived using Remote Sensing Technology – Rapport final – Projet de recherche Eurostat (non publié)

JAAKKOLA, O. (1998) : Multi-scale categorical data bases with automatic generalisation transformations based on map algebra – Cartography and geographic information systems, Vol. 25, N° 4, pages 195-207

JANSEN, L.J.M. ; DI GREGORIO A. (1999) : The Problems of Current Land Cover classifications: Development of a new approach. Dans Eurostat (1999) : Land Cover and Land Use information systems for European Union Policy Needs. Compte rendu du séminaire de Luxembourg les 21-23 janvier 1998

LAURINI, R. (1998) : Spatial multi-database topological continuity and indexing: a step towards seamless GIS data interoperability – International Journal on Geographic Information Science, Vol. 12, N°4 1998, pages 373-402

LILLESAND, T.M.; KIEFER, R.W. (1994) : Remote sensing and image interpretation. 3rd ed. Wiley, New York

LUND, HG. (1998) : A “Forest” by any other name. Paper in Land Use, Land Use change and forestry in the Kyoto protocol. Environmental Science & Policy, Elsevier Science publications

McCARTHY, D, (1996) : International Earth Rotation Service (IERS) Technical Note 21 – IERS Conventions, Observatoire de Paris

MÜCHER, C.A., STOMPH, T.J., Fresco L.O. (1993) : Proposal for a global land use classification FAO, Rome 37 pages

MEGRIN (1999) : Geographical Data Description Directory Version 2.1 - GDDD Data Model.- http://www.megrin.org/GDDD/Overview.html

NASA Goddard Space Flight Applied Information Sciences Branch Center (1999) : Remote Sensing and Photo Interpretation Tutorial, disponible à l’adresse http://rst.gsfc.nasa.gov/TofC/table.html

National Imagery and Mapping Agency (2000) : Technical Report 8350.2 3rd edition: World Geodetic System 1984. - Bethesda

OFFICE FÉDÉRAL DE LA STATISTIQUE (1996) : Statistique suisse de la superficie. L’utilisation des sols dans les cantons. Résultats par commune 1979/85 et 1992/97. Cette série (disponible en français et en allemand) donne les derniers résultats de l’enquête de suivi réalisée dans chaque canton. Les premiers volumes ont été publiés en 1996, et le dernier sortira probablement en 2001.

OLSON, J. M. (1998) : Maps as Representations of the World, NCGIA Core Curriculum in GIS, National Center for Geographic Information and Analysis, université de Californie, Santa Barbara, Unit 020, http://www.ncgia.ucsb.edu/giscc/units/u020/u020.html, sorti le 23 décembre 1998.

POIRET, M. (1997) : Groupe de travail sur Les Statistiques d’utilisation des sols

QUATTROCHI, D. A., SIU-NGAN LAM, Nina (1991) : Perspectives on Integrating Multiscale. Multitemporal Remote Sensing Data with Geographic Information Systems – The Integration of Remote Sensing and Geographic Information Systems – Compte rendu d’une session spéciale organisée à l’occasion de la convention annuelle Baltimore 1991 ACSM-ASPRS, ASPRS Bethesda Maryland USA

RADERMACHER, W. (1988) : Gedanken zu einer Basisklassifikation der Bodennutzungen

ROBINSON, A. ; MORRISON, J. ; MUEHRCKE, P. C. ; KIMERLING, J. et GUPTILL, S. C. (1995) : Elements of Cartography.- 6th Edition. New York: John Wiley and Sons, Inc.

SABINS, F.F. (1987) : Remote Sensing Principles and Interpretation - 2nd Edition, (W.H. Freeman & Co.), 449 pages

SIMPSON, G.G. (1961) : Principles of Animal Taxonomy, Columbia University Press, New York

SOKAL, R. (1974) : Classification: purposes, principles, progress, prospects dans Science vol. 185, n° 4197, pages 111-123

SNEATH, PHA. et SOKAL R. (1973) : Numerical taxonomy, Freeman and Co.

STATISTICS NORWAY (2000) : Land use statistics for urban settlements - Rapport final pour Eurostat

STEFANAKIS, E. SELLIS, T. (1998) : Enhancing operations with spatial access methods in a database management system for GIS - Cartography and Geographic Information Systems, Vol. 25 N°4, 1998, pages 16-32

STOTT, A. ; HAINES-YOUNG, R. (1996) : Linking land cover, intensity of use and botanical diversity in an accounting framework in the UK.

SUTCLIFFE, J.P. (1993) : Concept, class and category in the tradition of Aristotle. Chapter 3 in Van Mecheler et al. (eds.) Categories and concepts: theoretical views and inductive data analysis, Academic press, London

SYSAME (1992) : Cadastres fonciers nationaux - Synthèse de l'étude comparative des cadastres nationaux réalisée pour le casier viticole communautaire

TAYLOR, C. ; SANNIER, C. ; DELINCE, J. ; GALLEGO, F.J. (1997) : Regional Crop Inventories in Europe assisted by Remote Sensing: 1988-1993. Synthesis Report of the MARS Project - Action 1. Office des publications officielles des Communautés européennes, Luxembourg, EUR 17318

TOMPPO, E. (1992) : Multi-source National Forest Inventory of Finland - Finnish Forest Research Institute - Research Papers 444, Helsinki

University du Texas, Austin - Department of Geography (1999) : The Geographer's Craft Project - Austin (URL : http://www.utexas.edu/depts/grg/gcraft/about/about.html)

WEISSEL, J. et. al (1999) : Glossary – Lamont-Doherty Earth Observatory of Columbia University ; http://rst.gsfc.nasa.gov/Glossary/glossary.html

WÜSTER, E. (1971) : Les classifications de notions et de thèmes ; Différences essentielles et applications vol. 22 n°3 et 4, Bibliothèque Inforterm

WYATT, B. (1997) : Guidelines for Land Use and Land Cover description and classification, Institute of Terrestrial Ecology - Report to FAO